

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 15:25:00 ; Search time 97 Seconds
(without alignments)
3251.750 Million cell updates/sec

Title: US-09-671-687A-3
Perfect score: 5034
Sequence: 1 MSSGLWSQEKVTSYWEERI.....RLLCDAYMCYQSPMSLYK 949

Scoring table: BLOSOM62
Gapop 10.0 , Gapext 0.5

Searched: 1421835 seqs, 332370683 residues

Total number of hits satisfying chosen parameters: 1421835

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published Applications AA.*
- 1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/prodata/1/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/prodata/1/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/prodata/1/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/prodata/1/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/prodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/prodata/1/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/prodata/1/pubpaa/US09A_PUBCOMB.pep.*
 - 10: /cgn2_6/prodata/1/pubpaa/US09B_PUBCOMB.pep.*
 - 11: /cgn2_6/prodata/1/pubpaa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/prodata/1/pubpaa/US09_NEW_PUB.pep.*
 - 13: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/prodata/1/pubpaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/prodata/1/pubpaa/US10D_PUBCOMB.pep.*
 - 17: /cgn2_6/prodata/1/pubpaa/US10_NEW_PUB.pep.*
 - 18: /cgn2_6/prodata/1/pubpaa/US11_NEW_PUB.pep.*
 - 19: /cgn2_6/prodata/1/pubpaa/US60_NEW_PUB.pep.*
 - 20: /cgn2_6/prodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4983	99.0	960	16	US-10-788-792-250
2	4971.5	98.8	956	16	US-09-851-673-4
3	4971.5	98.8	956	16	US-10-755-889-490
4	3862	76.7	731	16	US-10-921-707-9
5	1275	25.3	261	15	US-10-264-237-1609
6	1173	23.3	238	15	US-10-264-237-1610
7	597	11.9	113	9	US-09-864-761-34675
8	146	2.9	354	13	US-10-087-192-120
9	143	2.8	547	15	US-10-108-260A-2751
10	143	2.8	547	15	US-10-275-595A-13
11	135.5	2.7	306	15	US-10-104-047-2990
12	135.5	2.7	721	11	US-09-764-875-857
13	135.5	2.7	2273	15	US-10-282-122A-66115

14	134.5	2.7	721	9	US-09-764-868-731	Sequence 731, Appl
15	134.5	2.7	721	11	US-09-764-875-1140	Sequence 1140, Ap
16	134.5	2.7	816	16	US-10-437-963-132799	Sequence 132799,
17	134	2.7	307	14	US-10-106-698-5606	Sequence 5606, Ap
18	133.5	2.7	1319	16	US-10-408-765A-343	Sequence 343, App
19	132.5	2.6	439	13	US-10-087-192-117	Sequence 117, App
20	130	2.6	717	15	US-10-369-493-22287	Sequence 22287, A
21	129.5	2.6	1281	16	US-10-363-829-373	Sequence 373, App
22	129.5	2.6	3298	14	US-10-160-758-16	Sequence 16, Appl
23	129.5	2.6	3298	14	US-10-174-677-8	Sequence 8, Appli
24	129.5	2.6	3298	15	US-10-120-801-51	Sequence 51, Appl
25	129.5	2.6	3298	15	US-10-210-172-50	Sequence 50, Appl
26	127.5	2.5	3217	15	US-10-311-623-8	Sequence 8, Appli
27	127	2.5	803	16	US-10-437-963-132702	Sequence 132702,
28	126.5	2.5	808	16	US-10-437-963-132681	Sequence 132681,
29	124	2.5	1474	14	US-10-225-567A-522	Sequence 522, App
30	124	2.5	1474	15	US-10-292-798-914	Sequence 914, App
31	124	2.5	2621	16	US-10-437-963-122168	Sequence 122168,
32	123.5	2.5	814	16	US-10-437-963-185098	Sequence 185098,
33	122.5	2.4	803	14	US-10-349-436-33	Sequence 33, Appl
34	122.5	2.4	892	15	US-10-276-774-1800	Sequence 1800, Ap
35	122.5	2.4	907	17	US-10-491-213-8	Sequence 8, Appli
36	122.5	2.4	1120	16	US-10-437-963-152821	Sequence 152821,
37	122.5	2.4	1282	16	US-10-437-963-110654	Sequence 110654,
38	122.5	2.4	1953	15	US-10-369-493-1945	Sequence 1945, Ap
39	122	2.4	2263	16	US-10-408-765A-2231	Sequence 2231, Ap
40	121	2.4	995	16	US-10-437-963-105117	Sequence 105117,
41	120.5	2.4	867	15	US-10-282-122A-57767	Sequence 57767, A
42	120.5	2.4	1033	9	US-09-888-615-75	Sequence 75, Appl
43	120.5	2.4	1109	15	US-10-425-114-72939	Sequence 72939, A
44	120.5	2.4	1471	10	US-09-998-027-4	Sequence 4, Appli
45	120.5	2.4	1471	14	US-10-165-099-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1

US-10-788-792-250
; Sequence 250, Application US/10788792
; Publication No. US20040191819A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Pharmaceuticals Corporation
; APPLICANT: Eveleigh, Deepa
; APPLICANT: Bigwood, Douglas
; TITLE OF INVENTION: EXPRESSION PROFILES FOR BREAST CANCER AND METHODS OF USE
; FILE REFERENCE: 5152
; CURRENT APPLICATION NUMBER: US/10788,792
; CURRENT FILING DATE: 2004-02-27
; PRIOR APPLICATION NUMBER: US 60/450,655
; PRIOR FILING DATE: 2003-02-28
; NUMBER OF SEQ ID NOS: 254
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 250
; LENGTH: 960
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-788-792-250

Query Match	99.0%	Score 4983	DB 16	Length 960
Best Local Similarity	99.5%	Pred. No. 0		
Mismatches	948	Conservative	0	Mismatches 1; Indels 4; Gaps 4;
Qy	1	MSSGLWSQEKVTSYWEERIFYLLOQSCVTDKOTKLLKVPKSGISGOYIDRSVGHRI	60	
Db	8	MSSGLWSQEKVTSYWEERIFYLLOQSCVTDKOTKLLKVPKSGISGOYIDRSVGHRI	67	
Qy	61	PSAKGKNQIGLKILEOPHAFVDE-DVVINEKFTELLAIITNCEERFSLFKNRRLS	119	
Db	68	PSAKGKNQIGLKILEOPHAFVDEKDVINEKFTELLAIITNCEERFSLFKNRRLS	127	
Qy	120	KGLQIDVCPVKQLRSGEKFPVGRFGRPELLARTVSGIPFGVLLLEEGSGQFTDGV	179	

Db 128 KGLQIDVGPVKVQLRSGBEKFPGVVRFRGPLLAERTVSGIFPGVLLBEGRCQGFTDGV 187
Qy 180 YQKQLFQCDDECG-FVALDKLELIEDDDTALSDYAGPDTMQVLPPLPINSRSLK 238
Db 188 YQKQLFQCDDECGFVALDKLELIEDDDTALSDYAGPDTMQVLPPLPINSRSLK 247
Qy 239 GETIESGTIVFCVLPKESLGYFVGVDMDNPIGNWDGRPDGV-LCSFACVESTILLHN 297
Db 248 GETIESGTIVFCVLPKESLGYFVGVDMDNPIGNWDGRPDGV-LCSFACVESTILLHN 307
Qy 298 DIIPESVTOERRPPKLAFAFMSRGVGDGSSSHNKPATGTSDDGNNR-RSELFTYLANGSSV 356
Db 308 DIIPESVTOERRPPKLAFAFMSRGVGDGSSSHNKPATGTSDDGNNRSELFTYLANGSSV 367
Qy 357 DSQPSQSKNTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSLTTENRHSFPFSL 416
Db 368 DSQPSQSKNTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSLTTENRHSFPFSL 427
Qy 417 TKMPNTNGSIGHSPLSLSAQSVMEELNTPAQVESPPPLAMPNGSHGLEVGSLAEVKNPP 476
Db 428 TKMPNTNGSIGHSPLSLSAQSVMEELNTPAQVESPPPLAMPNGSHGLEVGSLAEVKNPP 487
Qy 477 FYGVIRWIGOPPLNEVLAGELEDEACAGTGTFRGTRYFTCALKALFVKLSKCRPDS 536
Db 488 FYGVIRWIGOPPLNEVLAGELEDEACAGTGTFRGTRYFTCALKALFVKLSKCRPDS 547
Qy 537 RASLQPVNSQIRCNLSLAFGGVYLSEVVENTPPKMEKEGLEIMIGKKGIQGHYNSCYL 596
Db 548 RASLQPVNSQIRCNLSLAFGGVYLSEVVENTPPKMEKEGLEIMIGKKGIQGHYNSCYL 607
Qy 597 DSTLFCFLFAFSSVLDTVLLRPKEKNVVEYSETQELLRTTEIVNPLRIYGVVCATKIMKL 656
Db 608 DSTLFCFLFAFSSVLDTVLLRPKEKNVVEYSETQELLRTTEIVNPLRIYGVVCATKIMKL 667
Qy 657 KILEKVEAASGFTSEEDKPEEFNLILFHHILRVEPLLKIRSAQKQVQDCYFYQIFMEKNE 716
Db 668 KILEKVEAASGFTSEEDKPEEFNLILFHHILRVEPLLKIRSAQKQVQDCYFYQIFMEKNE 727
Qy 717 KVGVPITQQLLEWSFINSNLKFAEAPSCLLIQMPRGKDFKLPFKKIPPSLELNIITDLED 776
Db 728 KVGVPITQQLLEWSFINSNLKFAEAPSCLLIQMPRGKDFKLPFKKIPPSLELNIITDLED 787
Qy 777 TPRQCRICGGLAMYECRECYDDPDISAGKIKQFCCTCNTQVHLHPKLNHKNPNVSLPKD 836
Db 788 TPRQCRICGGLAMYECRECYDDPDISAGKIKQFCCTCNTQVHLHPKLNHKNPNVSLPKD 847
Qy 837 LPDWDNRHGICPCQNMEFLAVLCIETSHYVAFVKYKDDSAWLPFDSMADRDGGQGNFI 896
Db 848 LPDWDNRHGICPCQNMEFLAVLCIETSHYVAFVKYKDDSAWLPFDSMADRDGGQGNFI 907
Qy 897 PQVTPCPEVGEYLKMSLEDLHLSDSRRIQGCARLLCDAYMCMYQSPMTSLYK 949
Db 908 PQVTPCPEVGEYLKMSLEDLHLSDSRRIQGCARLLCDAYMCMYQSPMTSLYK 960

RESULT 2

US-09-851-673-4
; Sequence 4, Application US/09851673
; Publication No. US20030165985A1
; GENERAL INFORMATION:
; APPLICANT: Derry, Jonathan
; APPLICANT: Fenslow, William
; APPLICANT: Dougall, William
; TITLE OF INVENTION: SCREENING ASSAYS FOR AGONISTS OR ANTAGONISTS OF CD40 SIGNALING
; FILE REFERENCE: 3198
; CURRENT APPLICATION NUMBER: US/09/851, 673
; CURRENT FILING DATE: 2001-05-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 4
; LENGTH: 956
; TYPE: PR1
; ORGANISM: Homo sapiens

US-09-851-673-4

Query Match 98.8%; Score 4971.5; DB 10; Length 956;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 948; Conservative 0; Mismatches 1; Indels 7; Gaps 5;

Qy 1 MSSGLMSQEKVTSPPYWEERIFYLLLOECSTVDKQTKLLKVPKSGISQYIQDRSVGHSHRI 60
Db 1 MSSGLMSQEKVTSPPYWEERIFYLLLOECSTVDKQTKLLKVPKSGISQYIQDRSVGHSHRI 60
Qy 61 PSAGKKNQIIGLKILQPHAVLPVDE-DVVEINEKFTTELLAINTCBERESLFPKNRRLS 119
Db 61 PSAGKKNQIIGLKILQPHAVLPVDE-DVVEINEKFTTELLAINTCBERESLFPKNRRLS 120
Qy 120 KGLQIDVGPVKVQLRSGBEKFPGVVRFRGPLLAERTVSGIFPGVLLBEGRCQGFTDGV 179
Db 121 KGLQIDVGPVKVQLRSGBEKFPGVVRFRGPLLAERTVSGIFPGVLLBEGRCQGFTDGV 180
Qy 180 YQKQLFQCDDECG-FVALDKLELIEDDDTALSDYAGPDTMQVLPPLPINSRSLK 238
Db 181 YQKQLFQCDDECGFVALDKLELIEDDDTALSDYAGPDTMQVLPPLPINSRSLK 240
Qy 239 GETIESGTIVFCVLPKESLGYFVGVDMDNPIGNWDGRPDGV-LCSFACVESTILLHN 297
Db 241 GETIESGTIVFCVLPKESLGYFVGVDMDNPIGNWDGRPDGV-LCSFACVESTILLHN 300
Qy 298 DIIPESVTOERRPPKLAFAFMSRGVGDGSSSHNKPATGTSDDGNNR-RSELFTYLANG 353
Db 301 DIIPESVTOERRPPKLAFAFMSRGVGDGSSSHNKPATGTSDDGNNRSELFTYLANG 360
Qy 354 SSVDSQPSQSKNTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSLTTENRHSPL 413
Db 361 SSVDSQPSQSKNTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSLTTENRHSPL 420
Qy 414 FSLTKMPNTNGSIGHSPLSLSAQSVMEELNTPAQVESPPPLAMPNGSHGLEVGSLAEVKE 473
Db 421 FSLTKMPNTNGSIGHSPLSLSAQSVMEELNTPAQVESPPPLAMPNGSHGLEVGSLAEVKE 480
Qy 474 NPFFYGVIRWIGOPPLNEVLAGELEDEACAGTGTFRGTRYFTCALKALFVKLSKCR 533
Db 481 NPFFYGVIRWIGOPPLNEVLAGELEDEACAGTGTFRGTRYFTCALKALFVKLSKCR 540
Qy 534 PDSRFASLQPVNSQIRCNLSLAFGGVYLSEVVENTPPKMEKEGLEIMIGKKGIQGHYNS 593
Db 541 PDSRFASLQPVNSQIRCNLSLAFGGVYLSEVVENTPPKMEKEGLEIMIGKKGIQGHYNS 600
Qy 594 CYLDSTLFCFLFAFSSVLDTVLLRPKEKNVVEYSETQELLRTTEIVNPLRIYGVVCATKIM 653
Db 601 CYLDSTLFCFLFAFSSVLDTVLLRPKEKNVVEYSETQELLRTTEIVNPLRIYGVVCATKIM 660
Qy 654 KLRKILEKVEAASGFTSEEDKPEEFNLILFHHILRVEPLLKIRSAQKQVQDCYFYQIFME 713
Db 661 KLRKILEKVEAASGFTSEEDKPEEFNLILFHHILRVEPLLKIRSAQKQVQDCYFYQIFME 720
Qy 714 KNEKVGVPITQQLLEWSFINSNLKFAEAPSCLLIQMPRGKDFKLPFKKIPPSLELNIITDL 773
Db 721 KNEKVGVPITQQLLEWSFINSNLKFAEAPSCLLIQMPRGKDFKLPFKKIPPSLELNIITDL 780
Qy 774 LEDTPRQCRICGGLAMYECRECYDDPDISAGKIKQFCCTCNTQVHLHPKLNHKNPNVSL 833
Db 781 LEDTPRQCRICGGLAMYECRECYDDPDISAGKIKQFCCTCNTQVHLHPKLNHKNPNVSL 840
Qy 834 PKDLPDWDNRHGICPCQNMEFLAVLCIETSHYVAFVKYKDDSAWLPFDSMADRDGGQGN 893
Db 841 PKDLPDWDNRHGICPCQNMEFLAVLCIETSHYVAFVKYKDDSAWLPFDSMADRDGGQGN 900
Qy 894 FNIPOVTPCPEVGEYLKMSLEDLHLSDSRRIQGCARLLCDAYMCMYQSPMTSLYK 949
Db 901 FNIPOVTPCPEVGEYLKMSLEDLHLSDSRRIQGCARLLCDAYMCMYQSPMTSLYK 956

RESULT 3

US-10-755-889-490

Sequence 490, Application US/10755889
Publication No. US20040171823A1
GENERAL INFORMATION:
APPLICANT: Bristol-Myers Squibb Company
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-KB
TITLE OF INVENTION: PATHWAY
FILE REFERENCE: D0284 NP
CURRENT APPLICATION NUMBER: US/10/755,889
PRIOR FILING DATE: 2004-01-13
PRIOR APPLICATION NUMBER: U.S. 60/440,068
PRIOR FILING DATE: 2003-01-14
PRIOR APPLICATION NUMBER: U.S. 60/469,757
PRIOR FILING DATE: 2003-05-12
NUMBER OF SEQ ID NOS: 823
SOFTWARE: Patent in version 3.2
SEQ ID NO 490
LENGTH: 956
TYPE: PRT
ORGANISM: Homo sapiens
US-10-755-889-490

Query Match 98.8%; Score 4971.5; DB 16; Length 956;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 948; Conservative 0; Mismatches 1; Indels 7; Gaps 5;

QY 1 MSSGLWSQEKVTSPLYWEERIFVLLQECVTDKQTKLLKVPKGSIGQYIDRSVGHRI 60
DB 1 MSSGLWSQEKVTSPLYWEERIFVLLQECVTDKQTKLLKVPKGSIGQYIDRSVGHRI 60
QY 61 PSAGKKQIQIGKILFOPHVLVDDE-DVVEINEKFTTELLAITNCEERFSLFKNNRLS 119
DB 61 PSAGKKQIQIGKILFOPHVLVDDE-DVVEINEKFTTELLAITNCEERFSLFKNNRLS 120
QY 120 KGLQIDVGCVPVKQLRSBEKEFGVVRFRGPLLAERTVSGIFPGVELLEEGRGQGTG 179
DB 121 KGLQIDVGCVPVKQLRSBEKEFGVVRFRGPLLAERTVSGIFPGVELLEEGRGQGTG 180
QY 180 YQKQLFQDCEDCG-FVALDKLELIEDDDTALESYAGPGDMQVLPPELINSRVSLKG 238
DB 181 YQKQLFQDCEDCGFVVALDKLELIEDDDTALESYAGPGDMQVLPPELINSRVSLKV 240
QY 239 GTTIESGTIVFCVLPKGSLSGVFVGVDMNDPIGNWGRFDGV-LCSFACVESTILLHN 297
DB 241 GTTIESGTIVFCVLPKGSLSGVFVGVDMNDPIGNWGRFDGV-LCSFACVESTILLHN 300
QY 298 DIIP---ESVTQERRPPKLAFMSRGVGDGKSSSHNKPATGSTDPGNNR-RSELFTVLAG 353
DB 301 DIIPALSESVTQERRPPKLAFMSRGVGDGKSSSHNKPATGSTDPGNNRSELFTVLAG 360
QY 354 SSVDSQPSQSKNTWYIDEVAEDPAKSLTEISTDFDRSSPPPLQPPVNSLITENRPHSLP 413
DB 361 SSVDSQPSQSKNTWYIDEVAEDPAKSLTEISTDFDRSSPPPLQPPVNSLITENRPHSLP 420
QY 414 FSLTKMPNTNGSIGHSPLSLSAQSVMEELNAPVQESPPPLAMPNGSHGLVGSIAEVE 473
DB 421 FSLTKMPNTNGSIGHSPLSLSAQSVMEELNAPVQESPPPLAMPNGSHGLVGSIAEVE 480
QY 474 NPPFGVIRWIGQPPGLNEVLAGELEDEACAGCTDGTFRGTRYFTCALKKALFKVLKSCR 533
DB 481 NPPFGVIRWIGQPPGLNEVLAGELEDEACAGCTDGTFRGTRYFTCALKKALFKVLKSCR 540
QY 534 PDSRFASLOPVNSQIIBRCNSLAFGGVLSVVEENTPPKMEKEGLETWIGKKGIQGHYNS 593
DB 541 PDSRFASLOPVNSQIIBRCNSLAFGGVLSVVEENTPPKMEKEGLETWIGKKGIQGHYNS 600
QY 594 CYLDDTLFCLFAFSSVLDVTLARPKENNDVEYSETQELLRTIEVNPLRIYGVVCATKIM 653
DB 601 CYLDDTLFCLFAFSSVLDVTLARPKENNDVEYSETQELLRTIEVNPLRIYGVVCATKIM 660
QY 654 KLRKILEKVEAASGFTSEBKDPEEFNLIFHHILRVEPLLLKRSAGKQVQDCYFYQIFME 713
DB 661 KLRKILEKVEAASGFTSEBKDPEEFNLIFHHILRVEPLLLKRSAGKQVQDCYFYQIFME 720

QY 714 KNEKVGVPITQOLLEWSFINSLKFAEAPSCIIQMPRFCKDPKLPKKIIPPSLELNTDL 773
DB 721 KNEKVGVPITQOLLEWSFINSLKFAEAPSCIIQMPRFCKDPKLPKKIIPPSLELNTDL 780
QY 774 LEDTPQRCRICGGLAMYECECYDDDPDISAGKIKQFCKTQNTQVHLHPKRLNHNKYNPVS 833
DB 781 LEDTPQRCRICGGLAMYECECYDDDPDISAGKIKQFCKTQNTQVHLHPKRLNHNKYNPVS 840
QY 834 PKDLPDWRHGCIPQNNMELFAVLCTETSHYVAFVKYGHKDDSAWLFFDSMADRDGGONG 893
DB 841 PKDLPDWRHGCIPQNNMELFAVLCTETSHYVAFVKYGHKDDSAWLFFDSMADRDGGONG 900
QY 894 FNIPQVTPCPEVGEYLKMSLEDLSLDSRRIQCCARLLCDAYMCMYQSPMTSLYK 949
DB 901 FNIPQVTPCPEVGEYLKMSLEDLSLDSRRIQCCARLLCDAYMCMYQSPMTSLYK 956

RESULT 4
US-10-921-707-9
Sequence 9, Application US/10921707
Publication No. US20050003447A1
GENERAL INFORMATION:
APPLICANT: INCYTE PHARMACEUTICALS, INC.
APPLICANT: LAL, Preeti
APPLICANT: TANG, Y. Tom
APPLICANT: YUE, Henry
APPLICANT: HILLMAN, Jennifer L.
APPLICANT: BANDMAN, Olga
APPLICANT: CORLEY, Neil C.
APPLICANT: GUEGLER, Karl J.
APPLICANT: PATTERSON, Chandra
APPLICANT: AZIMZAI, Yalda
APPLICANT: BAUGHN, Mariah R.
TITLE OF INVENTION: HUMAN CYTOSKELETON ASSOCIATED PROTEINS
FILE REFERENCE: PF-0594 PCT
CURRENT APPLICATION NUMBER: US/10/921,707
CURRENT FILING DATE: 2004-08-19
PRIOR APPLICATION NUMBER: US/09/786,797
PRIOR FILING DATE: 2001-06-25
PRIOR APPLICATION NUMBER: 09/156,470; unassigned; 60/131,321
PRIOR FILING DATE: 1998-09-18; 1998-09-18; 1999-04-27
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PERL Program
SEQ ID NO 9
LENGTH: 731
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No: 2363327
US-10-921-707-9

Query Match 76.7%; Score 3862; DB 16; Length 731;
Best Local Similarity 99.8%; Pred. No. 6.1e-315;
Matches 728; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY 221 MQVELPPELINSRVSLKGGTIESGTIVFCVLPKGSLSGVFVGVDMNDPIGNWGRFDG 280
DB 1 MQVELPPELINSRVSLKGGTIESGTIVFCVLPKGSLSGVFVGVDMNDPIGNWGRFDG 60
QY 281 V-LCSFACVESTILLHNDIIIPESVTQERRPPKLAFMSRGVGDGSSSHNKPATGSTSD 339
DB 61 VQVCSFACVESTILLHNDIIIPESVTQERRPPKLAFMSRGVGDGSSSHNKPATGSTSD 120
QY 340 PGNNR-RSELFTVLAGSSVDSQPSQSKNTWYIDEVAEDPAKSLTEISTDFDRSSPPPLQPP 398
DB 121 PGNNRSELFTVLAGSSVDSQPSQSKNTWYIDEVAEDPAKSLTEISTDFDRSSPPPLQPP 180
QY 399 PVNSLTENRPHSLPFSLTAMPNTNGSIGHSPLSLSAQSVMEELNAPVQESPLAMPNG 458
DB 181 PVNSLTENRPHSLPFSLTAMPNTNGSIGHSPLSLSAQSVMEELNAPVQESPLAMPNG 240
QY 459 NSHGLEVGSIAEVEKPPFYGVIRWIGQPPGLNEVLAGELEDEACAGCTDGTFRGTRYFT 518

Abandoned Failure to Resp.
to an OA.
still last day
foreign priority

```
Db 241 NSHGLEVGLAEVKNPPYGVIRWIGQPPGLNEVLAGELEDEACAGTGTGTRGTRYFT 300
Qy 519 CALKALFVKLKS CRPDSRFSALQPVSNQIERCNSLAFGGYLSEVVEENTPPKMEKEGLE 578
Db 301 CALKALFVKLKS CRPDSRFSALQPVSNQIERCNSLAFGGYLSEVVEENTPPKMEKEGLE 360
Qy 579 IMIGKKGIQGHYNSCYLSTLFCPLPAPSSVLDTVLLRPKEKNDVVEYSETQELLATEIV 638
Db 361 IMIGKKGIQGHYNSCYLSTLFCPLPAPSSVLDTVLLRPKEKNDVVEYSETQELLATEIV 420
Qy 639 NPLRIYGVYCATKIMKRLKILEKVEAASGFTSEKDPESFLNLFPHILRVEPLLKIRSA 698
Db 421 NPLRIYGVYCATKIMKRLKILEKVEAASGFTSEKDPESFLNLFPHILRVEPLLKIRSA 480
Qy 699 GQKVQDVCYQIFMEKNEKVGVTPIQQLLEWFSFNSLNKFAEAPSCLIIQMPRFGKDFKL 758
Db 481 GQKVQDVCYQIFMEKNEKVGVTPIQQLLEWFSFNSLNKFAEAPSCLIIQMPRFGKDFKL 540
Qy 759 FKIPFSLNITDLEDTPQCRICGGLAMYECCYDDPDISAGKIKQFCKTCTNTQVH 818
Db 541 FKIPFSLNITDLEDTPQCRICGGLAMYECCYDDPDISAGKIKQFCKTCTNTQVH 600
Qy 819 LHPKRLNHNKYNPVS LKPLDPDWRHGCIPQNMELFAVLCTIETSHYVAFVKYKDDSAW 878
Db 601 LHPKRLNHNKYNPVS LKPLDPDWRHGCIPQNMELFAVLCTIETSHYVAFVKYKDDSAW 660
Qy 879 LFPDSMADRDGGGNGFNIQVPTPCPEVGEYKMSLEDLSRRTOGCARRLLCDAYMC 938
Db 661 LFPDSMADRDGGGNGFNIQVPTPCPEVGEYKMSLEDLSRRTOGCARRLLCDAYMC 720
Qy 939 MYQSPTMSLYK 949
Db 721 MYQSPTMSLYK 731

RESULT 5
US-10-264-237-1609
; Sequence 1609, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA131PI
; CURRENT APPLICATION NUMBER: US/10/264,237
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 1609
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (209)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (218)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (227)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (257)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
US-10-264-237-1609
```

```
Query Match 25.3%; Score 1275; DB 15; Length 261;
Best Local Similarity 98.7%; Pred. No. 2.6e-98;
Matches 234; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 653 MKLRKILEKVEAASGFTSEKDPESFLNLFPHILRVEPLLKIRSAQKQVQDCYFYQIFM 712
Db 1 MKLRKILEKVEAASGFTSEKDPESFLNLFPHILRVEPLLKIRSAQKQVQDCYFYQIFM 60
Qy 713 EKNEKVGVTPIQQLLEWFSFNSLNKFAEAPSCLIIQMPRFGKDFKLKIPFSLNITD 772
Db 61 EKNEKVGVTPIQQLLEWFSFNSLNKFAEAPSCLIIQMPRFGKDFKLKIPFSLNITD 120
Qy 773 LLEDTPQCRICGGLAMYECCYDDPDISAGKIKQFCKTCTNTQVHLHPKRLNHNKYNPVS 832
Db 121 LLEDTPQCRICGGLAMYECCYDDPDISAGKIKQFCKTCTNTQVHLHPKRLNHNKYNPVS 180
Qy 833 LKPLDPDWRHGCIPQNMELFAVLCTIETSHYVAFVKYKDDSAWLFPSMADRDG 889
Db 181 LKPLDPDWRHGCIPQNMELFAVLCTIETSHYVAFVKYKDDSAWLFPSMADRDG 237

RESULT 6
US-10-264-237-1610
; Sequence 1610, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA131PI
; CURRENT APPLICATION NUMBER: US/10/264,237
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 1610
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (38)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (187)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (227)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
US-10-264-237-1610

Query Match 23.3%; Score 1173; DB 15; Length 238;
Best Local Similarity 98.2%; Pred. No. 8.3e-90;
Matches 224; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 419 MPNTNGSIGHSPLSLSAQSVMEELNTPAVQESPPPLMPPGNSHGLEVGLAEVKNPPFY 478
Db 1 MPNTNGSIGHSPLSLSAQSVMEELNTPAVQESPPPLMPPGNSHGLEVGLAEVKNPPFY 60
Qy 479 GVIRWIGQPPGLNEVLAGELEDEACAGTGTGTRGTRYFTCALKKALFVKLKS CRPDSRF 538
Db 61 GVIRWIGQPPGLNEVLAGELEDEACAGTGTGTRGTRYFTCALKKALFVKLKS CRPDSRF 120
Qy 539 ASLQPVSNQIERCNSLAFGGYLSEVVEENTPPKMEKEGLEIMICKKGIQGHYNSCYLDS 598
Db 121 ASLQPVSNQIERCNSLAFGGYLSEVVEENTPPKMEKEGLEIMICKKGIQGHYNSCYLDS 180
Qy 599 TLFCLFAFSSVLDTVLLRPKEKNDVEYSETQELLRTEIVNPLRIYGY 646
```


Db 181 TLFLCFXFSVLDVLRPKKNDVEYSETQELLRTVNPURINXY 228

RESULT 7

US-09-864-761-34675
; Sequence 34675, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864, 761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263, 6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 34675
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007728.1
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.98
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.92
; OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HEL1, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.89
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4
; OTHER INFORMATION: EST HUMAN HIT: A1130924.1, EVALUE 5.00e-62
; OTHER INFORMATION: SWISSPROT HIT: Q03164, EVALUE 8.90e-01

US-09-864-761-34675

Query Match 11.9%; Score 597; DB 9; Length 113;
Best Local Similarity 100.0%; Pred. No. 6.1e-42;
Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 387 DFDSSPPLOPPVNSLTNRRFHSLLPFSLTMTKPNNTNGSICHSPLSLSAQSVMEELNTAP 446
Db 1 DFDSSPPLOPPVNSLTNRRFHSLLPFSLTMTKPNNTNGSICHSPLSLSAQSVMEELNTAP 60
QY 447 VQSPPLAMPNGSHGLEVGLAEVKNPPFYGVIRWIGQPPGLNEVLAGELEL 499
Db 61 VQSPPLAMPNGSHGLEVGLAEVKNPPFYGVIRWIGQPPGLNEVLAGELEL 113

RESULT 8

US-10-087-192-120
; Sequence 120, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 120
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-087-192-120

Query Match 2.9%; Score 146; DB 13; Length 354;
Best Local Similarity 20.4%; Pred. No. 0.0031;
Matches 89; Conservative 48; Mismatches 147; Indels 152; Gaps 19;

QY 219 DTMQVELPPLLEINSRVSLKG-----GETIESGTIVFCVDLPKGESLGYFVGVDMDN 269
Db 53 ETQEEFVDVDFRGVGVWVNGKPGFIQFLGET-----QFAPGQ-----WAGIVLDE 98
QY 270 PIGNWGRFDGVLCSFACVESTILLHINDIIPESVTOERRPPKLAFAWSRGVGDGSSSHN 329
Db 99 PIGNWGRFDGVLCSFACVESTILLHINDIIPESVTOERRPPKLAFAWSRGVGDGSSSHN 134
QY 330 KPKATGSTDGPNRRSELFTYLNSSVDSDQSKSKNTWYIDEVAEDPAKSLTEISTDFD 389
Db 135 EDEANGLQTTFASRATSPCTCTSTASWVSSSPSTPSN-----IPQKPSQFA----- 179
QY 390 RSSPPLQPPVNSLTNRRFHSLLPFSLTMTKPNNTNGSICHSPLSLSAQSVMEELNTAPQE 449
Db 180 -AKEPSATPISNLT-----KTASSEISNLSAGSIKK 211
QY 450 SPPLAMPNGSHGLEVGLAEVKNPPFYGVIRWIGQPPGLNEVLAGELEDEACGCTDG 509
Db 212 -----GERELKIGDRVLVGGTGA--GVVRFLGETDFAKGEWCGVEL--DEPLGKNDG 259
QY 510 TFRGTRYFTCALKKALFVKLKSCEPDRSFASLQPVNQIERCNSLAFGGLYSEVVEENTP 569
Db 260 AVAGTRYFQCPKYGLFA-----PVHKVTIGF-----PSTTP 292
QY 570 PKMEKGL-EIMIGKKKGIQGHVNSCYLDSTLFCFLFAFSSVLDTVLLRPKEKNDVEYYS 628
Db 293 AKANAVRRVMTTASLRKSPASSLS-----MSSVASSVSSRPS----- 335
QY 629 TQELLRTETVNPRIY 644

RESULT 11

Query Match 2.7%; Score 135.5; DB 15; Length 306;

RESULT 12

```

/ APPLICANT: AU, H.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: ELTRA.034A
/ CURRENT APPLICATION NUMBER: US/10/282,122A
/ CURRENT FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347

```



```
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1249
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1140
; LENGTH: 721
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (123)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-875-1140

Query Match      2.7%; Score 134.5; DB 11; Length 721;
Best Local Similarity 23.1%; Pred. No. 0.092;
Matches 82; Conservative 44; Mismatches 132; Indels 97; Gaps 14;

QY 193 GFVALDKLEIEDDD-----TALE-----SDVAGGDTMQVELPPL 228
Db 132 GYEDLDTFKLEEDDLDELNIRDPEHRAVLLTAVELLQYDSDSGQEKLVDSQGL 191
QY 229 E-INSRVS--LKGETIESGTIVFCVLPKESLGYFVGVDMDNPICNMDGREDGVLCSP 285
Db 192 SGCSPRDSCYESSLENLENGKTRKASLLSAKSTEPSLKSFERNQLGNY----- 240
QY 286 ACVESTILHINDIIPESVTOERRPPKLAFMMSGVGDKGSSSHNKPATGSTDGPNRRS 345
Db 241 ---PTLPLMKSGDALKQGEGR-----LGGGLAPDTSKCDPPGVTLNKL---NRRS 287
QY 346 -----ELFYTLNG-SSVDSQPSQSKNTWYID-----EVAEDPAKSLTEISTDFD 389
Db 288 LPVSIICRSCETLEGPTQVDTWPRSHSLDDLQVEPGAEDVPTETPPQIVPEVPQKT 347
QY 390 RSSPPLQPPPVNS-----LTTENRRFHSLPFSLTKMPTNGSIGHSPLSLSAQSVMEELN 443
Db 348 ASSTKAQPLEQDSAVDNALLTQSKFSEFPQKLT-TKKLEGSIAASGRGL----- 396
QY 444 TAPVQESPPPL-----AMPPGNHSHGLEVGSLAEVKENPPFYGVIRWIGQPPGLN 491
Db 397 -----SPPQCLPRNYDAQPPGAKHGLARTPLEGHRKKGHEFEGRHPLGTKEGVD 445
```

Search completed: April 18, 2005, 15:35:35
Job time : 100 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: April 18, 2005, 18:28:51 ; Search time 1146 Seconds
(without alignments)
5023.830 Million cell updates/sec

Title: US-09-671-687A-3
Perfect score: 5034
Sequence: 1 MSSGLWSQKVTSPWEERI.....RLLCDAYMCYQSPMTSLYK 949

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5622541 seqs, 303335566 residues
Total number of hits satisfying chosen parameters: 11245082

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgn2_1/USPTO_spool_p/US09671687/runat_18042005_115020_17247/app_query.fasta_1.1095
-DB=Published Applications NA - QPMF=fastap - SUFFIX=rnppb - MINMATCH=0.1
-LOOPCL=0 - LOOPEXT=0 - UNITS=bits - START=1 - END=-1 - MATRIX=blosum62
-TRANS=human40.cdi - LIST=45 - DOALIGN=200 - THR SCORE=pct - THR MAX=100
-THR MIN=0 - ALIGN=15 - MODES=LOCAL - OUTFMF=ptc - NORM=ext - HEADSIZE=500 - MINLEN=0
-MAXLEN=200000000 - USRR=US09671687 @CGN 1 1 723 runat_18042005_115020_17247
-NCPU=6 - ICPU=3 - NO MMAP - LARGEQUERY - NEG SCORES=0 - WAIT - DSPBLOCK=100
-LONGLOG - DEV_TIMEOUT=120 - WARN_TIMEOUT=30 - THREADS=1 - XGAPOP=10 - XGAPEXT=0.5
-FGAPOP=6 - FGAPEXT=7 - YGAPOP=10 - YGAPEXT=0.5 - DELOP=6 - DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:
19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	4971.5	98.8	5371	10	US-09-851-673-3	Sequence 3, Appli
2	4971.5	98.8	5371	18	US-10-755-889-489	Sequence 489, App
3	4942	98.2	4527	18	US-10-921-707-25	Sequence 25, Appl
4	3676	73.0	2523	15	US-10-037-270-290	Sequence 290, App
5	3676	73.0	2523	17	US-10-117-722-290	Sequence 3, Appli
6	2468	49.0	2116	18	US-10-761-370-3	Sequence 204, App
7	1778	35.3	1151	17	US-10-264-237-204	Sequence 205, App
8	1494	29.7	1013	17	US-10-264-237-205	Sequence 4253, Ap
9	1102	21.9	617	11	US-09-969-034-4253	Sequence 4254, Ap
10	879	17.5	557	11	US-09-969-034-4254	Sequence 1633, Ap
11	626	12.4	376	9	US-09-983-965-1633	Sequence 1845, A
12	545	10.8	309	9	US-09-864-761-18145	Sequence 1846, Ap
13	512	10.2	483	9	US-09-864-761-1386	Sequence 35878, A
14	412	8.2	425	10	US-09-918-995-35878	Sequence 1764, Ap
15	160.5	3.2	4847	17	US-10-152-319A-1764	Sequence 305, App
16	151	3.0	5857	10	US-09-873-367C-305	Sequence 3, Appli
17	148	2.9	3051	17	US-10-843-641A-305	Sequence 119, App
18	146	2.9	1380	13	US-10-087-192-119	Sequence 1545, Ap
19	146	2.9	5563	18	US-10-723-860-1545	Sequence 5373, Ap
20	146	2.9	5607	18	US-10-723-860-5973	Sequence 47, Appl
21	146	2.9	5607	18	US-10-723-860-5973	Sequence 308, App
22	143	2.8	3287	17	US-10-275-595A-47	Sequence 45974, A
23	143	2.8	3370	17	US-10-108-260A-308	Sequence 30316, A
24	143	2.8	3592	14	US-09-814-353-20078	Sequence 6166, Ap
25	138	2.7	2154	17	US-10-369-493-45974	Sequence 1020, Ap
26	138	2.7	2451	18	US-10-437-963-30316	Sequence 259, App
27	136.5	2.7	3409	18	US-10-723-860-6166	Sequence 25931, A
28	135.5	2.7	2618	17	US-10-104-047-1020	Sequence 56, Appl
29	135.5	2.7	3103	11	US-09-764-875-259	Sequence 118, App
30	135.5	2.7	6822	17	US-10-282-122A-29931	Sequence 542, App
31	135	2.7	2643	9	US-09-350-874-56	Sequence 1329, Ap
32	135	2.7	2643	15	US-10-106-989-56	Sequence 919, App
33	134.5	2.7	3103	9	US-09-764-868-118	Sequence 352, App
34	134.5	2.7	3103	11	US-09-764-875-542	Sequence 116, App
35	134	2.7	1099	15	US-10-106-698-1329	Sequence 65, Appl
36	134	2.7	2855	17	US-10-094-749-919	Sequence 450, App
37	133	2.6	3245	17	US-10-120-988-352	Sequence 121, Appl
38	132.5	2.6	1416	13	US-10-087-192-116	Sequence 10, Appl
39	132	2.6	3338	19	US-10-491-213-65	Sequence 49, Appl
40	130	2.6	2676	17	US-10-276-774-450	Sequence 58, Appl
41	129.5	2.6	4839	18	US-10-363-829-121	
42	129.5	2.6	10531	14	US-10-160-758-10	
43	129.5	2.6	10759	17	US-10-210-172-49	
44	128.5	2.6	2446	9	US-09-801-275-1	
45	128.5	2.6	2446	16	US-10-170-789-58	

ALIGNMENTS

RESULT 1
US-09-851-673-3
; Sequence 3, Application US/09851673
; Publication No. US20030165985A1
; GENERAL INFORMATION:
; APPLICANT: Derry, Jonathan
; APPLICANT: Fanslow, William
; APPLICANT: Dougall, William
; TITLE OF INVENTION: SCREENING ASSAYS FOR AGONISTS OR ANTAGONISTS OF CD40 SIGNALING
; FILE REFERENCE: 3198
; CURRENT APPLICATION NUMBER: US/09/851.673
; CURRENT FILING DATE: 2001-05-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 5371
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS

! LOCATION: (392)..(3262)
US-09-851-673-3

Alignment Scores:

Pred. No.: 0 Length: 5371
Score: 4971.50 Matches: 948
Percent Similarity: 99.16% Conservative: 0
Best Local Similarity: 99.16% Mismatches: 1
Query Match: 98.76% Indels: 7
DB: 10 Gaps: 5

US-09-671-687A-3 (1-949) x US-09-851-673-3 (1-5371)

```
QY 1 MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluGluArgIle 20
DB 392 ATGAGTTTCAGGCTTATGGAGCCCAAGAAAAGTCCATCTCCCTTACTGGGAAGCGGATT 451
QY 21 PheTyrLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLys 40
DB 452 TTTTACTGCTTCTTCAAGAAATGACGATTACAGACAGTATATCAAGATCGTTCTGGGGCATTTCAAGGATT 571
QY 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60
DB 512 GTACCAAGAGGAGTAGACAGTATATCAAGATCGTTCTGGGGCATTTCAAGGATT 571
QY 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGlnGlnProHisAla 80
DB 572 CTTCTGCAAGGCAAGAAAATCAGATTGGATTAATAATCTCAGAGCAACCTCATGCA 631
QY 81 ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeu 99
DB 632 GTTCTCTTTGATGAAAAAGGATGTTGTAGAGATAAATGAAAAGTTTCACAGAGTTACTT 691
QY 100 LeuAlaIleThrAsnGlyGlnGluArgPheSerLeuPheLysAsnArgLeuSer 119
DB 692 TTGGCAATTTACCAATTTGTGAGGAGAGGTTGAGGCTCTTTAAAAACAGAAACAGACTAAGT 751
QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGlu 139
DB 752 AAGGGCTCCAAATAGACGTGGGCTCTCTGTAAGATGACAGTGAATCTGGGGAGAA 811
QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
DB 812 AAATTTCTGGAGTTGACGTTTCAGAGACCCCTGTAGCAGAGGACAGTCTCCGA 871
QY 160 IlePhePheGlyValGluLeuLeuGluGluArgGlyGlnGlyPheThrAspGlyVal 179
DB 872 ATATTTCTTTGGAGTTGAATTTGTGGAAGAAGGTCGTGGTCAAGGTTTCACTGACGGGGTG 931
QY 180 TyrGlnGlyGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
DB 932 TACCAAGGGAACACAGCTTTTTCAGTGTGATGAAGATTGTGGCGTGTGTTGTGATTTGGAC 991
QY 199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
DB 992 AAGCTAGAACTCTAGAGATGATGACCTGCATTTGGAAAGTATTTACGCGAGTCTCTGG 1051
QY 219 AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly 238
DB 1052 GACACAATGCGAGTCGAATCTCTCTCTTTGGAAATAAATCCAGAGTTTCTTTGAGGTT 1111
QY 239 GlyGluThrIleGlnSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
DB 1112 GGAAGAACAAATAGAAATCTGGAACAGTTATATCTGTGATGTTTGGCAGGAAAAGAAC 1171
QY 259 LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe 278
DB 1172 TTAGGATATTTGTTGGTGGACATGGAATACCTTATTGGCACTGGGATGGAAGATT 1231
QY 279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
DB 1232 GATGGAGTGCAGCTTTGTAGTTTTCGTGTGTTGAAAGTACAAATCTTATTGCACATCAAT 1291
```

```
QY 298 AspIleIlePro-----GluSerValThrGlnGluArgProProLysLeuAla 314
DB 1292 GATATCATCCAGCTTTATCAGAGAGTGTGACCGAGAAAGGAGGCTCCCAACTTGGC 1351
QY 315 PheMetSerArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThr 334
DB 1352 TTTATGTCAAGAGGTGTTGGGCAAAAGGTTTCCAGTCTATTAATAAACCAAGGCTTACA 1411
QY 335 GlySerThrSerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGly 353
DB 1412 GGATCTACCTCAGACCCCTGGAATAGAACAGATCTGAATATTTTATACCTTTAAATGG 1471
QY 354 SerSerValAspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGluVal 373
DB 1472 TCTTCTGTTGACTCACAAACACCAATCCAAATCAAAAAATACATGCTACATTGATGAAGTT 1531
QY 374 AlaGluAspProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerPro 393
DB 1532 GCAGAAAGACCTCGCAAAATCTCTTACAGAGATATCTACAGACTTTGACCGTTCTTACCA 1591
QY 394 ProLeuGlnProProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuPro 413
DB 1592 CCACTCCAGCCTCTCTCTGTGAACCTCCTGACCCAGCAACAGATTCACATCTTTTACCA 1651
QY 414 PheSerLeuThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeu 433
DB 1652 TTCAGTCTCACCAAGATGCCAATATCAATGGAAGTATTGGCCACAGTCCACATTTCTCTG 1711
QY 434 SerAlaGlnSerValMetGluGluLeuAsnThrAlaProValGlnGlnSerProProLeu 453
DB 1712 TCAGCCAGCTGTGTAATGGAAGAGCTAAACACCTGCACCCCTCCAAGAGAGTCCACCCCTG 1771
QY 454 AlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGlu 473
DB 1772 GCCATCCCTCTGGGAACCTCACATGCTGTAGAAAGTGGGCTCATTTGGCTGAAGTTAAGAG 1831
QY 474 AsnProProPheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluVal 493
DB 1832 AACCTCTCTTCTATGGGGTAATCCGTTGGATCCGTCAGCCAGCCAGGACTGAATGAAGTG 1891
QY 494 LeuAlaGlyLeuGluLeuAspGluCysAlaGlyCysThrAspGlyThrPheArgGly 513
DB 1892 CTGCTGGACTGGAATCGGAAGATGAGTGTGCAAGGCTGTACGGATGGAACCTTCAGAGGC 1951
QY 514 ThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPheValLysLysSerCysArg 533
DB 1952 ACTCGGTATTTCACTTCCCTGTCCTGGAAGAGCGCTGTTGTGAAACTGGAAGAGCTGAGG 2011
QY 534 ProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSer 553
DB 2012 CCTGACTCTAGGTTTGCACTATTTGCAAGCCGCTTTCCAAATCAGATTGAGCGCTGTAACCT 2071
QY 554 LeuAlaPheGlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGlu 573
DB 2072 TTAGCATTTGGAGGCTACTTAAGTGAAGTAGTAGAAGAAAATACTCCACCAAAAAATGGA 2131
QY 574 LysGluGlyLeuGluLeuMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSer 593
DB 2132 AAAGAAGGCTTGGAGATTAATGATTGGGAAGAAAGGATCCAGGGTCAATTACAATCT 2191
QY 594 CysTyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrVal 613
DB 2192 TGTACTTAGACTCAACCTTATTCTGCTTATTTGCTTTTATGTTCTGTCTGGACACTGTG 2251
QY 614 LeuLeuArgProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeu 633
DB 2252 TTACTTAGACCCAAAGAAAGACGATGATAGATATTTATAGTGAACCCAGAGACTACTG 2311
QY 634 ArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMet 653
DB 2312 AGGACAGAAATTTGTAATCTCTCAGAAATATATGATATGTGTGTGTCACAAAAAATTATG 2371
QY 654 LysLeuArgLysIleLeuGluLysValGluAlaSerGlyPheThrSerGluGluLys 673
```



```
Db 2372 AAACCTGAGGAAATACCTTGAAAGGTGGAGCTGCATCAGGATTTACCTCTGAAGAAAA 2431
Qy 674 AspProGluGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeu 693
Db 2432 GATCCTGAGGAATCTTGAATATCTGTTTCATCATATTTTAAGGTGAGAACCTTTGCTA 2491
Qy 694 LysIleArgSerAlaGlyGlnIysValGlnAspCysTyrPheTyrGlnIlePheMetGlu 713
Db 2492 AAAATAAGATCAGCAGGTCAAAAGGTACAAAGATTTGTTACTTCTATCAAAATTTTATGGAA 2551
Qy 714 LysAsnGluLysValGlyValProThrIleGlnGlnLeuGluTrpSerPheIleAsn 733
Db 2552 AAAATGAGAAAGTTGGCTTCCCAAAATTCAGCAGTTGTTAGAATGGTCTTTTATCAAC 2611
Qy 734 SerAsnLeuLysPheAlaGluAlaProSerCysLeuIleGlnMetProArgPheGly 753
Db 2612 AGTAACCTGAATTTGCAGAGGCACCATCATGTCTGATATTACAGATGCTTCGATTTGGA 2671
Qy 754 LysAspPheLysLeuPheLysIlePheProSerLeuGluLeuAsnIleThrAspLeu 773
Db 2672 AAAGACTTTAAACTATTAAAAAAATTTTCTCTCTCTGGAATTAATAATACAGATTTA 2731
Qy 774 LeuGluAspThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArg 793
Db 2732 CTTGAAGACACTCCCAAGAGTCCCGATATGTGGAGGGCTTGCAATGTATGAGTGTAGA 2791
Qy 794 GluCysTyrAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCys 813
Db 2792 GAATGCTACGAGCATCCGACATCTCAGCTGGAAAAAATCAAGCAGTTTGTAAACCTGC 2851
Qy 814 AsnThrGlnValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeu 833
Db 2852 AACACTCAAGTCCACCTTCATCCGAGAGGCTGAATCATAAATATAACCCAGTGTCACTT 2911
Qy 834 ProLysAspLeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGlu 853
Db 2912 CCNAAAGACTTACCCGACTGGGACTGGAGACACGGCTGCATCCCTGGCCAGAAATATGGAG 2971
Qy 854 LeuPheAlaValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLys 873
Db 2972 TTATTGTGCTTCTCTGCATAGAAACAAGCACATATGTGCTTTTGTGAAGTATGGGAAG 3031
Qy 874 AspAspSerAlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGly 893
Db 3032 GACGATTCGCTTGCTCTCTTTGACAGATCGCCGATCGGGATGCTGTCAGAAATGGC 3091
Qy 894 PheAsnIleProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeu 913
Db 3092 TTCAACATTTCTCAAGTCACTCCCATGCCAGAGAGTAGGAGATCTTGAAGATGTCTCTG 3151
Qy 914 GluAspLeuHisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCys 933
Db 3152 GAAGACCTGCATTTCTTGGACTCCAGAGAATCCAAAGGCTGTGCACGAAGACTGCTTTGT 3211
Qy 934 AspAlaTyrMetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
Db 3212 GATGCATATATGTGCATGTACCAAGAGTCCAAACAATGAGTTGTACAAA 3259
```

RESULT 2

```
US-10-755-889-489
; Sequence 489, Application US/10755889
; Publication No. US20040171823A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-kB
; TITLE OF INVENTION: PATHWAY
; FILE REFERENCE: D0284 NP
; CURRENT APPLICATION NUMBER: US/10/755,889
; CURRENT FILING DATE: 2004-01-13
; PRIOR APPLICATION NUMBER: U.S. 60/440,068
; PRIOR FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: U.S. 60/469,757
```

```
; PRIOR FILING DATE: 2003-05-12
; NUMBER OF SEQ ID NOS: 823
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 489
; LENGTH: 5371
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-755-889-489

Alignment Scores:
Pred. No.: 0 Length: 5371
Score: 4971.50 Matches: 948
Percent Similarity: 99.16% Conservative: 0
Best Local Similarity: 99.16% Mismatches: 1
Query Match: 98.76% Indels: 7
DB: 18 Gaps: 5

US-09-671-687A-3 (1-949) x US-10-755-889-489 (1-5371)
```

```
Qy 1 MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluGluArgIle 20
Db 392 ATGAGTTTCAGGCTTATGGAGCCAAAGAAAAGTCACTTCACTTACCTTACCTGGAAGAGCGGATT 451
Qy 21 PheTyrIleLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLeuLys 40
Db 452 TTTTACTTGTCTTCTCAAGAAATGCAGCGTTACAGACAAACAAACACAAAGCTCCTTAAA 511
Qy 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60
Db 512 GTACCGAAGGAAGATATAGACAGATATATTCAGATCGTTCTGTGGGCATTCAGGATT 571
Qy 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla 80
Db 572 CCTTCTCAAAAGGCAAGAAAATACATTCGATTGGAATTAATAATTTCTAGACCAACCTCATGCA 631
Qy 81 ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu 99
Db 632 GTTCTCTTTCTGATGAAAAGGATGTTGTAGAGATAAAATGAAAAGTTTACAGAGTTACTT 691
Qy 100 LeuAlaIleThrAsnCysGluGluAtgPheSerLeuPheLysAsnAtgAsnArgLeuSer 119
Db 692 TTGGCAATTAACCAATTGTGAGGAGAGGTTGAGCTGTTTAAATAACAGAAAACAGACTAAGT 751
Qy 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuAtgSerGlyGluGlu 139
Db 752 AAAGGCTCCAAATAGACGTGGGCTGCTCTGTGAAGTACAGCTGAGATCTGGGGGAAGAA 811
Qy 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
Db 812 AAATTTCTCGAGTTGTACGCTTCAGAGAGCCCCCTGTTAGCAGAGAGACAGTCTCCGGA 871
Qy 160 IlePhePheGlyValGluLeuLeuGluGluGlyArgGlyGlnGlyPheThrAspGlyVal 179
Db 872 ATATTCTTTGAGTTGAATTTGCTGGAAGAGGTCGTGTCGAAGTTTTCACCTGACGGGTG 931
Qy 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
Db 932 TACCAAGGGAACACAGCTTTTTCAGTGTGATGAAGATTGTGGCGCTGTTGTTGTCATTTGAC 991
Qy 199 LysLeuLeuLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
Db 992 AAGCTAGAACTCATAGAAAGATGATGACACTGCAATGGAAAGTATTACGAGGCTCTGGG 1051
Qy 219 AspThrMetGlnValGluLeuProProLeuGluLeuIleAsnSerArgValSerLeuLysGly 238
Db 1052 GACCAATGCAAGGTGCAAACTTCTCTCTTTGGAATAAATCAACTCCAGAGTTTCTTTGAAGTT 1111
Qy 239 GlyClnThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
Db 1112 GGAGAAAACAATAGAAATCTGGAACAGTTATATCTGTGATGTTTGTCCAGGAAAAAGAAC 1171
Qy 259 LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe 278
```

Db 1172 TTAGGATATTTCTTGGTGGACATGGATAACCCCTATTGGCAACTGGGATGGAAGATT 1231
QY 279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
Db 1232 GATGGAGTGCAGCTTTGTAGTTTGGCTGTGTGAAAGTACAATTCTATTGGACATCAAT 1291
QY 298 AsnIleIlePro-----GluSerValThrGlnGluArgArgProProLysLeuAla 314
Db 1292 GATATCATCCAGCTTTATCAGAGAGTGTGACGACGAAAGGAGGCTCCCAAACTTGGC 1351
QY 315 PheMetSerArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThr 334
Db 1352 TTTATGTCAAGAGGTGTGGGCAAAAGTTTCATCCAGTCATATAAACAAGGCTACA 1411
QY 335 GlySerThrSerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGly 353
Db 1412 GGATCTACCTCAGACCCCTGGAATAAGAACAGATCTGAATTAATTTATACCTTAATGGG 1471
QY 354 SerSerValAspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGluVal 373
Db 1472 TCTTCTGTGTGACTCACACCAATCCAAATCAAAATAATACATGGTACATTTGATGAAGTT 1531
QY 374 AlaGluAspProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerPro 393
Db 1532 GCAGAAAGCCCTGCAAAATCTCTTACAGAGATATCTACAGACTTTTGACCCGTTCTTCAACA 1591
QY 394 ProLeuGlnProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuPro 413
Db 1592 CCACCTCCAGCCCTCTCTGTGAATCACTGACCCAGACAGATCCACCTCTTTACCA 1651
QY 414 PheSerLeuThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeu 433
Db 1652 TTCACTCTCACCAAGATGCCAATACCAATGGAAGTATTGGCCACAGTCCACTTCTCTG 1711
QY 434 SerAlaGlnSerValMetGluGluLeuAsnThrAlaProValGlnGluSerProLeu 453
Db 1712 TCAGCCCACTGTGTAAATGGAAGAGCTAAACACTGCACCCGTCACAGAGAGTCCACCCCTG 1771
QY 454 AlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGlu 473
Db 1772 GCCATGCCCTCTGGGAATCACTCATGTCTTAGAAGTGGGCTCATTTGGCTGAAGTTAAGGAG 1831
QY 474 AsnProProPheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluVal 493
Db 1832 AACCCCTCTTCTATGGGTAAATCCGTTGGATCGGTGAGCCACGAGCTGAATGAAGTG 1891
QY 494 LeuAlaGlyLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGly 513
Db 1892 CTCGCTGGACTGGAACTGGAAAGATGAGTGTGCAGGCTGTACGGATGGAACCTTCAGAGGC 1951
QY 514 ThrArgTyrPheThrCysAlaLeuLysAlaLeuPheValLysLeuLysSerCysArg 533
Db 1952 ACTCGGTATTTCACCTGTGCCCTTGAAGAGGCGCTGTGTGAACTGAAGAGCTGAGG 2011
QY 534 ProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSer 553
Db 2012 CCTGACTCTAGGTTTGATCATTTGACGCGGTTTCCAAATCAGATTGAGCGCTTAACCTCT 2071
QY 554 LeuAlaPheGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGlu 573
Db 2072 TTAGCATTTGGAGGCTACTTAAGTAGAGTAGTAGAAGAAATACTCCACCAAAATGGAA 2131
QY 574 LysGluGlyLeuLurleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSer 593
Db 2132 AAAGAAGGCTTGGAGATAATGATTGGGAAGAAAGAGGCAATCCAGGCTCAATTCAATCT 2191
QY 594 CysTyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrVal 613
Db 2192 TGTACTTAGACTCAACCTTATTCTGCTTATTGCTTTTAGTTCTGTTCTGGACACTGTG 2251
QY 614 LeuLeuArgProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeu 633
Db 2252 TTACTTTAGACCCCAAGAAAGACGATGTAGATAATTATAGTGAACCCCAAGAGCTACTG 2311

QY 634 ArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMet 653
Db 2312 AGCACAGAAATGTTAATCTCTGAGATATATGATATGTTGTGTCACAAAAATATG 2371
QY 654 LysLeuArgLysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLys 673
Db 2372 AAACCTGAGAAATACTTGAAGAGTGGAGGCTGCATCAGGATTTACCTCTGAAGAAAA 2431
QY 674 AspProGluGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeu 693
Db 2432 GATCTGAGGAATTTCTTGAATATTCTGTTTCATCATATTTTAAGGGTAGAACCTTTGCTA 2491
QY 694 LysIleArgSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGlu 713
Db 2492 AAAATTAAGATCAGCAGGTCAAAAGGTACAAGATTGTTACTTCTATCAAAATTTTATGGAA 2551
QY 714 LysAsnGluLysValGlyValProThrIleGlnGlnLeuLeuGluTrpSerPheIleAsn 733
Db 2552 AAAAATGAGAAAGTTGGCGTTCCACAAATTCACAGCTGTGTAGAATGGTCTTTTATCAAC 2611
QY 734 SerAsnLeuLysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGly 753
Db 2612 AGTAACCTGAATTTTCAGAGGACCATCATGTCATATTTCAGATGCCCTCGATTGGA 2671
QY 754 LysAspPheLysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeu 773
Db 2672 AAAGACTTTAAACTATTAAAAAAATTTTCTCTCTGGAATTAATAACAGATTTA 2731
QY 774 LeuGluAspThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArg 793
Db 2732 CTTGAAGACACTCCACAGACGTGCCGATATGTGGAGGCTTGCATATGTATGAGTGTAGA 2791
QY 794 GluCysTyrAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCys 813
Db 2792 GAATGCTAGACCATCCGACATCTCAGCTGGAAAAATCAAGCAGTTTGTGAAACCTGC 2851
QY 814 AsnThrGlnValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeu 833
Db 2852 AACACTCAAGTCCACCTTCATCCGAGAGGCTGAATCATATAATAACCCAGTGTCACTT 2911
QY 834 ProLysAspLeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGlu 853
Db 2912 CCAAAAGACTTACCCGACTGGAGCTGGAGACACGCTGCATCCCTGCCAGAAATATGGAG 2971
QY 854 LeuPheAlaValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLys 873
Db 2972 TTATTTGCTGTCTCTGCATAGAAACAAGCCACTATGTTGCTTTTGTGAAGTATGGGAAG 3031
QY 874 AspAspSerAlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGly 893
Db 3032 GACGATTTGCTGGCTGCTTCTTTTGACGATGCGCGATCGGATGGTGTGAGATGGC 3091
QY 894 PheAsnIleProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeu 913
Db 3092 TTTCAACATTTCTCAAGTCACTCCCATGCCAAGAGTAGGAGACTTTGAAGATGTCTCTG 3151
QY 914 GluAspLeuHisSerLeuAspSerArgArgIleGlnGlyCysAlaArgLeuLeuCys 933
Db 3152 GAAGACCTGCAATTTCTTGGACTCCAGGAGAAATCCAAGGCTGTGCACCAAGACTCTTGT 3211
QY 934 AspAlaTyrMetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
Db 3212 GATGCATATATGTCATGTACACAGATCCAAACAATGAGTTGTACAAA 3259

RESULT 3

US-10-921-707-25
; Sequence 25, Application US/10921707
; Publication No. US20050003447A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE PHARMACEUTICALS, INC.
; APPLICANT: LAL, Preeti
; APPLICANT: TANG, Y. Tom

APPLICANT: YUE, Henry
 APPLICANT: HILLMAN, Jennifer L.
 APPLICANT: BANDMAN, Olga
 APPLICANT: CORLEY, Neil C.
 APPLICANT: GUEGLER, Karl J.
 APPLICANT: PATTERSON, Chandra
 APPLICANT: AZIMZAI, Yalda
 APPLICANT: BAUGHN, Mariah R.
 TITLE OF INVENTION: HUMAN CYTOSKELETON ASSOCIATED PROTEINS
 FILE REFERENCE: PF-0594 PCT
 CURRENT APPLICATION NUMBER: US/10/921,707
 CURRENT FILING DATE: 2004-08-19
 PRIOR APPLICATION NUMBER: US/09/786,797 ← *SEA 25 = here but this = 371 filed*
 PRIOR FILING DATE: 2001-06-25
 PRIOR APPLICATION NUMBER: 09/156,470, unassigned; 60/131,321
 PRIOR FILING DATE: 1998-09-18; 1998-09-18; 1999-04-27
 NUMBER OF SEQ ID NOS: 32
 SOFTWARE: PERL Program
 SEQ ID NO 25
 LENGTH: 4527
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: misc feature
 OTHER INFORMATION: Incyte ID No: 2363327
 US-10-921-707-25

not here
SEA ID 21 in this provisional.

Alignment Scores:
 Pred. No.: 0 Length: 4527
 Score: 4942.00 Matches: 944
 Percent Similarity: 99.37% Conservations: 0
 Best Local Similarity: 99.37% Mismatches: 2
 Query Match: 98.17% Indels: 5
 DB: 18 Gaps: 4

US-09-671-687A-3 (1-949) x US-10-921-707-25 (1-4527)

Qy	4	GlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluGluArgIlePheTyrLeu	23
Db	270	GGCTTATGAGCCAGAAAAAGTCATCTCACCCCTACTGGGAGAGCGGATTTTACTTG	329
Qy	24	LeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLeuLysValProLys	43
Db	330	CTTCTTCAAGATGCGACGGTTACAGACAAACAAACAAAGCTCTTAAAGTACCGAA-	388
Qy	44	GlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIleProSerAla	63
Db	389	GGAGTATAGGACAGTATATTCAGATCGTTCTGTGGGGCAITTCAGGATTCCTCTGCA	448
Qy	64	LysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAlaValLeuPhe	83
Db	449	AAAGGCAAGAAAAATCAGATTGGATTAAAAATTCAGAGCAACCTCATGCAATTCTCTT	508
Qy	84	ValAspGlu---AspValValGluLysAsnGluLysPheThrGluLeuLeuAlaIle	102
Db	509	GTTGATGAAAGGATCTGTAGAGATAAATGAAAGTTACAGAGTTACTTTTGGCAATT	568
Qy	103	ThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSerLysGlyLeu	122
Db	569	ACCAATTGTGAGAGAGGTTGACCCCTGTTTAAAAACAGAAACAGACTAAGTAAAGGCCCTC	628
Qy	123	GlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGluLysPhePro	142
Db	629	CAATAGACGTGGGCTGTCTGTGAAAGTACAGCTGAGATCTGGGGAAGAAAAATTTCTCT	688
Qy	143	GlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGlyIlePhePhe	162
Db	689	GGAGTTGTACGCTTTCAGAGACCCCTGTTAGCAGAGAGGACACTCTCCGGAATATCTTT	748
Qy	163	GlyValGluLeuLeuGluGluGlyArgGlyGlnGlyPheThrAspGlyValTyrGlnGly	182
Db	749	GGAGTTGAATTTGCGAAGAAAGTCGTGGTCAAGGTTTCTACCTGACGGGGTGTAACCAAGG	808

Qy	183	LysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAspLysLeuGlu	201
Db	809	AAACAGCTTTTTCAGTGTGATGAAGATTGGCGCTGTTTGTTCATTGGAACAAGCTAGAA	868
Qy	202	LeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGlyAspThrMet	221
Db	869	CTCATAGAAGATGATGACACTGCACTGGAAGTATGATACGCGAGGTCTCTGGGACACAATG	928
Qy	222	GlnValGluLeuProLeuGluIleAsnSerArgValSerLeuLysGlyGlyThr	241
Db	929	CAGGTGGAACCTCTCTCTCTGGAATAAATCCACAGAGTTCTTTGAAGGTTTGGAGAAACA	988
Qy	242	IleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSerLeuGlyTyr	261
Db	989	ATAGAATCTCGAAGCAGTTATTTCTGTGATGTTTTGCCAGGAAAGAAAGCTTAGGATAT	1048
Qy	262	PheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPheAspGlyVal	281
Db	1049	TTTGTGGTGGACATGGATAACCTATTGGCACTGGGATGGAGATTTGATGGAGTG	1108
Qy	282	---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsnAspIleIle	300
Db	1109	CAGCTTTGTAGTTTGGTGTGTTGAAAGTACAATTCTATTGCACATCAATGATATCATC	1168
Qy	301	ProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMetSerArgGlyVal	320
Db	1169	CCAGAGAGTGTGACGCGAGGAAAGGAGCCCTCCCAAACTTGCCTTTATGTCAGAGGTGT	1228
Qy	321	GlyAspLysGlySerSerHisAsnLysProLysAlaThrGlySerThrSerAspPro	340
Db	1229	GGGCAACAAAGTTTCATCCAGTCATAATAAACCAAGGCTACAGGATCTTACCTCAGACCCT	1288
Qy	341	GlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerValAspSerGln	359
Db	1289	GGAAATAGAAACAGACTGTAATTTTATACCTTAAATGGGTCTTCTGTGTGACTCACA	1348
Qy	360	ProGlnSerLysSerLysAsnThrTrpTyrIleAspGluValAlaGluAspProAlaLys	379
Db	1349	CCACAATCCAAATCAAAAAATACATGATGATGATGATGATGATGATGATGATGATGATG	1408
Qy	380	SerLeuThrGluIleSerThrAspPheAspArgSerProProLeuGlnProProPro	399
Db	1409	TCCTTTACAGAGATATCTACAGACTTTGACCGTTCCTTCCACCACCTCCACCTCTCTCT	1468
Qy	400	ValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeuThrLysMet	419
Db	1469	GTGAACCTCAGTACCCAGGAGACAGATTCACCTCTTTACCATTCAGTCTCACCAGATG	1528
Qy	420	ProAsnThrAsnGlySerIleGlyHisSerProLeuSerSerAlaGlnSerValMet	439
Db	1529	CCCAATACCAATGGAAGTATTGGCCACAGTCCACTTTCTCTGTGAGCCAGCTCTGTAATG	1588
Qy	440	GluGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetProProGlyAsn	459
Db	1589	GAAGAGCTAAACATGACCCCGTCCAAAGAGAGTCCACCTTTGGCCATGCTCTCTCTCT	1648
Qy	460	SerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPheTyrGly	479
Db	1649	TCATCTGCTAGAGTGGGCTCATTTGGCTGAAAGTAAAGGAAACCTCTCTCTCTCTCT	1708
Qy	480	ValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGlyLeuGluLeu	499
Db	1709	GTAAATCCGTTGGATCGGTGAGCCACACAGGACTGAATGAAGTGTCTCGTGGACTGGAAC	1768
Qy	500	GluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyrPheThrCys	519
Db	1769	GAAGATGAGTGTGAGGCTGTACGATGGAACCTTTCAGAGGCACTCGGTATTTTACCTGT	1828
Qy	520	AlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSerArgPheAla	539
Db	1829	GCCTTGAAGAAAGCGCTGTTTGTGAAACTGAAGAGCTGACGCGCTGACTCTAGGTTTGA	1888
Qy	540	SerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPheGlyTyr	559

```
Db 1889 TCATTGACGGGTTTCCAAATCAGATTGACGCTGTAACCTTTTAGCATTTGGAGCTAC 1948
Qy 560 LeuSerGluValValGluGluAanThrProProLysMetGluLysGluGlyLeuGluIle 579
Db 1949 TTAAGTGAAGTAGTAGAAGAAATATCTCCACCAAAATGGAAGAGAGGCTTGGAGATA 2008
Qy 580 MetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeuAspSerThr 599
Db 2009 ATGATTGGGAAGAAAGCATCCAGGGGTCAATACAAATCTCTGTACTTAGACTCAACC 2068
Qy 600 LeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArgProLysGlu 619
Db 2069 TTATTCTGCTATTCTGCTTTAGTTCTGTTCTGGACACTGTGTACTTAGACCCAAAGAA 2128
Qy 620 LysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGluIleValAsn 639
Db 2129 AAGAAGCATGTAGAATATTATAGTGAACCCAGAGACTACTGAGGACAGAAATGTTAAT 2188
Qy 640 ProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArgLysIleLeu 659
Db 2189 CCTCTGAGAATATATGATATGTGTGCCACAAAATATGAAACTGAGGAAAATACTT 2248
Qy 660 GluLysValGluAlaLaserGlyPheThrSerGluLysAspProGluGluPheLeu 679
Db 2249 GAAAAGGTGGAGGCTGCATCAGGATTTACCTCTGAAGAAAAGATCCTGAGGAATCTTG 2308
Qy 680 AsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArgSerAlaGly 699
Db 2309 AATATTCTGTTTCATCATATTTTAAAGGGTAGAACCTTTGCTAAAATAAGATCAGCAGGT 2368
Qy 700 GlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGluLysValGly 719
Db 2369 CAAAAGGTACAAGATTGTTACTTCTATCAATTTTATGGAATAATGAGAAAGTTGGC 2428
Qy 720 ValProThrIleGlnGlnLeuLeuGluThrSerPheIleAsnSerAsnLeuLysPheAla 739
Db 2429 GTTCCCAACAATTCAGCAGTTGTTAGAATGGTCTTTTATCAACAGTAACCTGAAATTTGCA 2488
Qy 740 GluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPheLysLeuPhe 759
Db 2489 GAGGCACCATCATGCTGTGATTATTCAGATGCCCTCGAATTTGGAAGAGACTTTAAACATTT 2548
Qy 760 LysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAspThrProArg 779
Db 2549 AAAAAATTTTCTCTCTGGATTAATATAACAGATTACTTGNAGACATCCCGA 2608
Qy 780 GlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyrAspAspPro 799
Db 2609 CAGTCCGGATATGTGGAGGCTTGCAATGTATGAGTGTAGAAATGCTACGACGATCCG 2668
Qy 800 AspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGlnValHisLeu 819
Db 2669 GACATCTCAGCTGGAAAAATCAAGCAGTTTGTAAAACTGCAACACTCAAGTCCACCTT 2728
Qy 820 HisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAspLeuProAsp 839
Db 2729 CATCCGAGAGGCTGAATCAATAATATAACCCAGTGTCACTTCCCAAGACTTACCCGAC 2788
Qy 840 TrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAlaValLeuCys 859
Db 2789 TGGGACTGGAGACACGCGCTGCATCCCTTGCAGAAATATGAGTATTATTGCTGTTCTCGC 2848
Qy 860 IleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSerAlaTrpLeu 879
Db 2849 ATAGAAACAAGCCACTATGTTGCTTTTGTGAAGTATGGGAAGGACGATTTGCTGCTGCTC 2908
Qy 880 PhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIleProGlnVal 899
Db 2909 TTCTTTGACAGCATGGCCGATGGGATGGGTGGTGCAGAAATGGCTTCAACATTCTCTCAAGTC 2968
Qy 900 ThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLeuHisSerLeu 919
```

```
Db 2969 ACCCATGCCAGAGTAGAGAGTACTTGAAGATGCTCTGGAAGACCTGCATCTCTTG 3028
Qy 920 AspSerArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyrMetCysMet 939
Db 3029 GACTCCAGGAGAAATCCAAAGCTGTGCACGAAGACTGCTTTGTGATGCATATATGTG 3088
Qy 940 TyrGlnSerProThrMetSerLeuTyrLys 949
Db 3089 TACCAGAGTCCAAACAATGAGTTTGTACAAA 3118
RESULT 4
US-10-037-270-290
; Sequence 290, Application US/10037270
; Publication No. US20030104529A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunding
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Tillinghast, John
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/10/037, 270
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt FL_genes Version 1.0
; SEQ ID NO 290
; LENGTH: 2523
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (81)..(2138)
US-10-037-270-290
Alignment Scores:
Pred. No.: 0 Length: 2523
Score: 3676.00 Matches: 692
Percent Similarity: 99.28% Conservative: 0
Best Local Similarity: 99.28% Mismatches: 3
Query Match: 73.02% Indels: 2
DB: 15 Gaps: 2
US-09-671-687A-3 (1-949) x US-10-037-270-290 (1-2523)
Qy 255 GlyLysGluSerLeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrp 274
Db 45 GGAATAATAAGAGCTTTTATATTTTGTGGTGTGACATGATGAACCCCTATTGGCAACTGG 104
Qy 275 AspGlyArgPheAspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeu 293
Db 105 GATGGAAGATTTGATGGAGTGCAGCTTTGTAGTTTGGTGTGTGAAAGTACAACTTA 164
Qy 294 LeuHisIleAsnAspIleIleProGluSerValThrGlnGluArgArgProProLysLeu 313
Db 165 TTGCACATCAATGATATCATCCAGAGAGTGTGACGAGGAAGAGGAGGCTCCCAAACTT 224
```

314 AlaPheMetSerArgGlyValGlyAspLysGlySerSerSerHisAsnLysProLysAla 333
225 GCCTTTATGTCAAGAGGTGTGGGCAAGGTTCTCCAGTCATATAAACAACCAAGGCT 284
334 ThrGlySerThrSerAspProGlyAsnArg- --ArgSerGluLeuPheTyrThrLeuAsn 352
285 ACAGGATCTACCTCAGACCCCTGAATAGAACAGATCTGAATATTTATATACCTTAAT 344
353 GlySerSerValAspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGlu 372
345 GGGTCTCTCTGTTGACTCACAACCAATCCAAATCAAAATAATACATGATGATGAA 404
373 ValAlaGluAspProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSer 392
405 GTTGCAAGAGACCCCTCAAAATCTCTTACAGAGATATCTACAGACTTTGACCGCTCTTCA 464
393 ProProLeuGlnProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeu 412
465 CCACCACTCCAGGCTCTCTGTGAATCTACTGACCAACGAGAACAGATTCACACTCTTTA 524
413 ProPheSerLeuThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSer 432
525 CCATTGAGTCTCACCAGATGCCAATACCAATGGAAGTATTGGCCACAGCTCCACTTCT 584
433 LeuSerAlaGlnSerValMetGluGluLeuAsnThrAlaProValGlnGluSerProPro 452
585 CTGTGAGCCAGCTCTGTAATGGAAGAGCTAAACACTGCACCCCTCCAAAGAGTCCACCC 644
453 LeuAlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLys 472
645 TTGGCCATCCCTCTGGGAACCTCACATGGTGTAGAAGTGGGCTCATTTGGCTGAAGTTAAG 704
473 GluAsnProProPheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGlu 492
705 GAGAACCTCTCTTCTATGGGTAACTCCGTGGATCGGTGAGCCACCAAGACTGAATGAA 764
493 ValLeuAlaGlyLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArg 512
765 GTGCTCGCTGGAGTGAATGGAAGATGAGTGTGCAGGCTGTACGGATGGAACCTTCAGA 824
513 GlyThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCys 532
825 GGCACCTCGGTATTTCACTGTGCCCTTGAAGAGCGCTGTTTGTGAACCTGAAGAGCTGC 884
533 ArgProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsn 552
885 AGCCCTGACTCTAGGTTTGCATCTTGCAGCCGCTTCCATCAGATTGAGCCCTGTAAC 944
553 SerLeuAlaPheGlyGlyTyrLeuSerGluValValGluLeuAsnThrProProLysMet 572
945 TCTTTAGCATTTGGAGGCTACTTTAAGTGAAGTAGTAGAAGAAATACTCCACCAAAATG 1004
573 GluLysGluGlyLeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsn 592
1005 GAAAAAGAGGCTTGGAGATAATGATTGGGAGAGAAAGAGGATCCAGGGTCAATTACAAT 1064
593 SerCysTyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerValLeuAspThr 612
1065 TCTTGTACTTAGACTCAACCTTATTCTGCTTATTTGCTTTTATGTTCTGTCTGGACACT 1124
613 ValLeuLeuArgProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeu 632
1125 GTGTTACTTAGACCCCAAGAAAGACGATGTAGATATATTATAGTGAACCCCAAGAGCTA 1184
633 LeuArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIle 652
1185 CTGAGGACAGAAATGTGTAATCTCTGAGATATATGATATGATGTGTGTCACAAAAATT 1244
653 MetLysLeuArgLysIleLeuGluLysValGluAlaSerGlyPheThrSerGluGlu 672
1245 ATGAAACTGAGGAAAAATCTTGAAGAGGTGGAGGCTGCATCAGGATTTTACCTCTGAAGA 1304

673 LysAspProGluGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeu 692
1305 AAAGATCCTCAGGAATTCITGAATATTCTGTTTCATCATATTTTAAAGGTTAGAACCTTGT 1364
693 LeuLysIleArgSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMet 712
1365 CTAAATAAAGATCAGCAGGTCAAAAGGTACAAGATTGTTACTTCTATCAAAATTTTATG 1424
713 GluLysAsnGluLysValGlyValProThrIleGlnGlnLeuLeuGluTyrSerPheIle 732
1425 GAAAAAATGAGAAGTTGGCGTTCGCCAATTCAGCAGTTGTTAGAAATGCTCTTTTATC 1484
733 AsnSerAsnLeuLysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPhe 752
1485 AACAGTAACCTGAAATTTGCAGAGGCACCATCATGTCGTGATTATTTCAGATGCTCGATTT 1544
753 GlyLysAspPheLysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAsp 772
1545 GGAAGAAGACTTTAAACTATTATTAATAAAATTTTCTCTCTGGAATTAATAATACAGAT 1604
773 LeuLeuGluAspThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCys 792
1605 TTACTTTGAAGACACTCCAGACAGTGCAGGATATGTGAGGCGCTTGCAATGTATGAGTGT 1664
793 ArgGluCysTyrAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThr 812
1665 AGAAGATGCTTACAGCAGTCCGACATCTCAGCTGGAAAAATCAAGCAGATTTTGTAACACC 1724
813 CysAsnThrGlnValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSer 832
1725 TGCAACACTCAGTCCACCTTCATCCAGAGGCTGAATCATATAATATAACCCAGTGTA 1784
833 LeuProLysAspLeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMet 852
1785 CTTCCCAAGACTTACCCGACTGGGACTGGAGACACGCGTCGATCCCTTGGCAGAATATG 1844
853 GluLeuPheAlaValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGly 872
1845 GAGTATTTCGTCTCTCATAGAAACAAGCCACTATGTTGCTTTTGTGAAGTATGGG 1904
873 LysAspAspSerAlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsn 892
1905 AAGCAGATTCGCTGGCTGCTTCTTTTGACAGCATGCGGATCGGATGGTGGTCAGAAT 1964
893 GlyPheAsnIleProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSer 912
1965 GGCTTCAACATTCCTCAAGTCACCCCATGCCAGAGTAGGAGAGTACTTGAAGATGCT 2024
913 LeuGluAspLeuHisSerLeuAspSerArgArgIleGlnGlyCysAlaArgLeuLeu 932
2025 CTGGAAGACCTGCATTCCTTGGACTCCAGAGAAATCCAAAGGCTGTCACGAAGACTGCTT 2084
933 CysAspAlaTyrMetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
2085 TGTGATGCATATATGTGATGTACCATGTCAGAGTCCCAATGAGTGTGTGACAA 2135

RESULT 5

US-10-117-722-290
; Sequence 290, Application US/10117722
; Publication No. US20030219744A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030219744A1el Nucleic Acids and
; TITLE OF INVENTION: Polydeptides
; FILE REFERENCE: 784CJP25CIP
; CURRENT APPLICATION NUMBER: US/10/117,722
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 09/620,312
; PRIOR FILING DATE: 2000-07-19

```

; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 290_
; LENGTH: 2523
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (81)..(2138)
US-10-117-722-290
Alignment Scores:
Pred. No.: 0 Length: 2523
Score: 3676.00 Matches: 692
Percent Similarity: 99.28% Conservative: 0
Best Local Similarity: 99.28% Mismatches: 3
Query Match: 73.02% Indels: 2
DB: 17 Gaps: 2

US-09-671-687A-3 (1-949) x US-10-117-722-290 (1-2523)
QY 255 GlyLysGluSerLeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrp 274
DB 45 GGAATAATAAGCTTTTATATTTTGGTGTGGACATGATGAACCTATTGGCAACTGG 104
QY 275 AspGlyArgPheAspGlyVal----LeuCysSerPheAlaCysValGluSerThrIleLeu 293
DB 105 GATGGGAAGATTTCATGAGGTGCGAGCTTTGTAGTTTTCGCGTGTGTGAAAGTACAACTTA 164
QY 294 LeuHisIleAsnAspIleIleProGluSerValThrGlnGluArgArgProProLysLeu 313
DB 165 TTCACATCAATGATATATCCAGAGAGGTGACCGAGAAAGGAGGCCCTCCCAACTT 224
QY 314 AlaPheMetSerArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAla 333
DB 225 GCCTTTATGTCAAGAGGTGTGGGACAAAGTTTCATCCAGTCAATAAACCAGGCT 284
QY 334 ThrGlySerThrSerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsn 352
DB 285 ACAGGATCTACCTCAGACCCTGGAAATAGAAACAGATCTGAATTAATTTATACCTTAAAT 344
QY 353 GlySerSerValAspSerGlnProGlnSerLysSerLysSerThrTrpTyrIleAspGlu 372
DB 345 GGCTCTCTGTGACTCACAAACCACCAATCCAAATCAAAATAATACATGGTACATTGATGAA 404
QY 373 ValAlaGluAspProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSer 392
DB 405 GTTGCAGAGACCCCTGCATAATCTCTTACAGATATCTACAGACTTTGGACCGTCTTCA 464
QY 393 ProProLeuGlnProProValAsnSerLeuThrGluAsnArgPheHisSerLeu 412
DB 465 CCACCACTCCAGCCTCCTCTGTAACCTCACTACCCACCCAGAACAGATTCCTTTA 524
QY 413 ProPheSerLeuThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSer 432
DB 525 CCATTTCAGTCTCCCAAGATGCCCAATACCAATGGGAAGTATTGGCCACAGTCCACTTTCT 584
QY 433 LeuSerAlaGlnSerValMetGluLeuAsnThrAlaProValGlnGluSerProPro 452
DB 585 CTGTACCCCGAGTCTGTAATGGAGAGACTAAACACTGCACCCCGCTCCAGAGAGTCCACCC 644
QY 453 LeuAlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLys 472
DB 645 TTGGCCATGCTCTCTGGGAACCTCACATGGTCTAGAGAGTGGGCTCATTTGGCTGAAGTTAAG 704
QY 473 GluAsnProProPheTyrGlyIleArgTrpIleGlyGlnProProGlyLeuAsnGlu 492
DB 705 GAGAACCCCTCTTCTATGGGTAATCCGTTGGATCGGTGAGCCAGCCAGGACTGAATGAA 764

```

```

QY 493 ValLeuAlaGlyLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArg 512
DB 765 GTGCTCGCTGGACTGGAACCTGGAAAGATGAGTGTGCAGGCTGTACGATGGAACCTTCAGA 824
QY 513 GlyThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCys 532
DB 825 GGCACCTCGGTATTTTACCTGTGCCCTGAGAGGCGCTGTTGTGAAACTGAGAGAGTGC 884
QY 533 ArgProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsn 552
DB 885 AGGCTGTAGCTTAGGTTTGCATCATTTGCAGCGGTTTCCCAATCAGATTGAGCGCTGTAAC 944
QY 553 SerLeuAlaPheGlyTyrLeuSerGluValValGluGluAsnThrProProLysMet 572
DB 945 TCTTTAGCATTTTGGAGGCTACTTAAGTGAAGTAGTAGAAGAAATACTCCACCAAAATG 1004
QY 573 GluLysGluGlyLeuGluIleMetIleGlyLysLysLysGlyIleGlnGlyHisTyrAsn 592
DB 1005 GAAAAAGAGAGGCTTGGAGATAATGATTGGGAAGAAAGAGCATCCAGGGTTCATTACAT 1064
QY 593 SerCysTyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThr 612
DB 1065 TCTTGTACTTAGACTCAACTTATTCTGCTTATTGCTTTTAGTCTGTTCTGGACACT 1124
QY 613 ValLeuLeuArgProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeu 632
DB 1125 GTGTTACTTAGACCCCAAGAAAGAACGATGTAGAATATTATATAGTGAACCCAGAGCTA 1184
QY 633 LeuArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIle 652
DB 1185 CTGAGACAGAAATTTGTAATCTCTGAGAAATATATGATATGTGTGTGCGCACAAATTT 1244
QY 653 MetLysLeuArgLysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGlu 672
DB 1245 ATGAAACTGAGGAATAATCTTGAAGAGTGGAGGCTGCATCAGGATTTACTCTGAGAA 1304
QY 673 LysAspProGluGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeu 692
DB 1305 AAAGATCCTGAGGAAATCTTGAATATTCTGTTTTCATCATATTATTAAGGTAAGACCTTG 1364
QY 693 LeuLysIleArgSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMet 712
DB 1365 CTAAAAATTAAGATCAGCAGGTCAAAAGGTACAAAGATTGTTACTTCTATCAAAATTTTATG 1424
QY 713 GluLysAsnGluLysValGlyValProThrIleGlnGlnLeuLeuGluTyrSerPheIle 732
DB 1425 GAAAAAATGAGAAAGTTGGCGTTCCCAATTTCCAGAGTTGTAGAAATGGTCTTTATC 1484
QY 733 AsnSerAsnLeuLysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPhe 752
DB 1485 AACAGTAACCTGAAATTTGCAGAGGCCACCATCATGCTGATTATTAGATGCCTCGATTT 1544
QY 753 GlyLysAspPheLysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAsp 772
DB 1545 GGAAGAGACTTTAAACTATTATAAAAAATTTTCTCTCTGGAATTTAAATATATACAGAT 1604
QY 773 LeuLeuGluAspThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCys 792
DB 1605 TTACTTGAAGACACTCCACAGAGTCCGATATGTGGAGGGCTTGCAGATGTATGAGTGT 1664
QY 793 ArgGluCysTyrAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThr 812
DB 1665 AGAGATGCTACCGATCCGACATCTCAGTCCGAAAAATCAAGCAGTTTGTGTAACACC 1724
QY 813 CysAsnThrGlnValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSer 832
DB 1725 TGCAACACTCAAGTCCACCTTCATCCGAGAGGCTGAATCATATAATATAACCCAGTGTCA 1784
QY 833 LeuProLysAspLeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMet 852
DB 1785 CTTCCCAAGACTTTACCCGACTGGAGTGGAGACACCGGCTGCATCCCTTGGCCAGAAATG 1844
QY 853 GluLeuPheAlaValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGly 872

```


Db 1845 GAGTTATTGCTGTTCTCTGCATAGAAACAAGCCATATGTTGCTTTTGTGAAGTATGGG 1904
QY LysAspAspSerAlaTrpLeuPheAspSerMetAlaAspArgAspGlyGlnAsn 892
Db 1905 AAGGACGATTCTGCTGGCTCTCTTTGACAGCATGGCCGATCGGATGCTGTCAGAAAT 1964
QY GlyPheAsnIleProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSer 912
Db 1965 GCGTTCAACATTCTCTCAAGTCAACCCCATGCCAAGAGTAGGAGACTTGAAGATGTCT 2024
QY LeuGluAspLeuHisSerLeuAspSerArgAtqIleGlnGlyCysAlaArgArgLeuLeu 932
Db 2025 CTGGAAGACTGCATTCTTGGACTCCAGGAGAAATCAAGGCTGTGCACGAAGACTGCTT 2084
QY CysAspAlaTyrMetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
Db 2085 TGTGATGCATATATGTGCATGTACCAGAGTCCAAACAATGAGTTGTACAAA 2135

RESULT 6

US-10-761-370-3
; Sequence 3, Application US/10761370
; Publication No. US20040219615A1
; GENERAL INFORMATION:
; APPLICANT: WALLACH, David
; APPLICANT: KOVALENKO, Andrei
; TITLE OF INVENTION: MODULATORS OF THE FUNCTION OF RECEPTORS OF THE TNF/NGF RECEPTOR R
; TITLE OF INVENTION: AND OTHER PROTEINS
; FILE REFERENCE: WALLACH=27
; CURRENT APPLICATION NUMBER: US/10/761,370
; CURRENT FILING DATE: 2004-01-22
; PRIOR APPLICATION NUMBER: US/09/646,403
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: IL 123758
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: PCT/IL99/00158
; PRIOR FILING DATE: 1999-03-18
; PRIOR APPLICATION NUMBER: IL 126024
; PRIOR FILING DATE: 1998-09-01
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 2116
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (691)..(691)
; OTHER INFORMATION: n is unknown.
US-10-761-370-3

Alignment Scores:
Pred. No.: 5,34e-261 Length: 2116
Score: 2468.00 Matches: 524
Percent Similarity: 91.21% Conservative: 5
Best Local Similarity: 90.34% Mismatches: 33
Query Match: 49.03% Indels: 22
DB: 18 Gaps: 6

US-09-671-687A-3 (1-949) x US-10-761-370-3 (1-2116)

QY 387 AspPheAspArgSerProLeuGlnProProValAsnSerLeuThrThrGlu 406
Db 15 GACTTTGACCGTCTTCCACCACCATCCAGCCCTCTCTGTGAACCTACGACACCCGAG 74
QY 407 AsnArgPheHisSerLeuProPheSerLeuThrLysMetProAsnThrAsnGlySerIle 426
Db 75 AACAGATTCCACTCTTTACCATTCAGTCTCACCAAGATGCCAATACCATTGGAAGTATT 134
QY 427 GlyHisSerProLeuSerLeuSerAlaGlnSerValMetGluLeuLeuAsnThrAlaPro 446
Db 135 GGCACAGTCCACTTCTCTCTGTCAGCCAGCTCTGTAATGAAGAGCTAAACACTGCAACC 194

QY 447 ValGlnGluSerProProLeuAlaMetProProGlyAsnSerHisGlyLeuGluValGly 466
Db 195 GTCACAGAGAGTCCACCTTTGGCCATGCCCTCTGGGAACCTCACATGGTCTAGAAAGTGGGC 254
QY 467 SerLeuAlaGluValLysGluAsnProProPheTyrGlyValIleArgTrpIleGlyGln 486
Db 255 TCATTGGCTGAAGCTTAAGGAGAACCCCTCTCTTCTATGGGGTAAATCCGTTGGATCGGTCTCAG 314
QY 487 ProProGlyLeuAsnGluValLeuAlaGlyLeuGluLeuGluAspGluCysAlaGlyCys 506
Db 315 CCACAGGAGTGAATGAAGTCTGCTGGACTGGAACTGGAAAGATGAGTGTGCGAGGCTGT 374
QY 507 ThrAspGlyThrPheArgGlyThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPhe 526
Db 375 ACCGATGGAACTTCAGAGGCACCTCGGTATTTACCTGTGCCCTCGAAGAGCGCGTGTITT 434
QY 527 ValLysLeuLysSerCysArgProAspSerArgPheAlaSerLeuGlnProValSerAsn 546
Db 435 GTGAACCTGAAGAGCTGCAGCCCTGACTCTAGGTTTGCATCATTTGCAGCCCGTTTCCAAT 494
QY 547 Gln-IleGluArgCysAsnSerLeuAlaPheGlyGlyTyrIleuSerGluValValGluG 566
Db 495 CAAGATTGAGCGCTGAACCTCTTAGCATTTGGAGGCTACTTAAAGTGAAGTAGT-GAAGA 553
QY 566 uAsnThrProProLysMetGluLysGluLeuLysMetIleGly---LysLysLys 585
Db 554 AAATACTCCACCAAAATGAAAGAAAGCTTGGAGATAATGATTTGGGGAAGAAAGAAA 613
QY 585 sGlyIleGlnGlyHisTyrAsnSerCysTyrLeuAspSerThrLeuPhe-CysLeuPhe- 604
Db 614 GGCATCCAGGGCTCATTACAAATTTCTGKTACTTAGACTCAACCTTATTCTTGCTTATTTK 673
QY 605 AlaPheSerSerValLeu-AspThr---ValLeuLeuArgProLysGluLysAsnAsp-- 622
Db 674 GCTTTTAGTTCCTGTCINGACACTGGTGTACTTCTAGACCCCAAGAAAGAAACGAT 733
QY 623 -ValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGluIleValAsnProLeuAr 642
Db 734 GTTAGAATATTTWTWKGMMACCCCAAGAGCTACTGAGGACAGAAATTTGTAATCTCTCGAG 793
QY 642 gIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArgLysIleLeuGluLysVa 662
Db 794 AATATATGGATATGTGTGCCCAAAATTTATGAAACTGAGGAAATTTACTTTGAAAGGT 853
QY 662 lGluAlaAlaSerGlyPheThrSerGluGluLysAspProGluGluPheLeuAsnIleLe 682
Db 854 GGAGGCTGCATCAGGATTTACCTCTGAAGAAAGATCTCTGAGGAATTTCTTGAATATTCT 913
QY 682 uPheHisHisIleLeuArgValGluProLeuLeuLysIleArgSerAlaGlyGlnLysVa 702
Db 914 GTTTCATCATATTTTAAGGGTAGAACCTTTGCTAAAAATAAGATCAGCAGGTCAAAAGGT 973
QY 702 lGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGluLysValGlyValProTh 722
Db 974 ACAAGATTTGTTACTTCTATCAAAATTTTATGGAAAAAATGAGAAAGTTGGCGTTCCAC 1033
QY 722 rIleGlnGlnLeuLeuGluTrpSerPheIleAsnSerAsnLeuLysPheAlaGluAlaPr 742
Db 1034 AATTACAGCAGTTGTTAGAATGGTCTTTTATCAACAGTAACCTGAAATTTGCAGAGGCACC 1093
QY 742 oSerCysLeuIleIleGlnMetProArgPheGlyLysAspPheLysLeuPheLysIle 762
Db 1094 ATCATGTCGATTTATTCAGATGCTCGATTTGGAAGAGACTTTAAACTATTATTAATAATT 1153
QY 762 ePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAspThrProArgGlnCysAr 782
Db 1154 -TTTCTCTCTCTGGAATTAGATATAACAGATTTACTTTGAAGACAC-CCCAGACAGTGGCG 1211
QY 782 gIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyrAspAspProAspIleSe 802
Db 1212 GATATGTGGAGGGCTTCAATGTATGATGATGATAA-GAATGCTACGACGATCGGACACCCAG 1270
QY 802 rAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGlnValHisLeuHisProly 822


```

Db      1271 C---TCGAAAACACGAGCTTTTGTAAACCTGCAACACTCAAGTCACCTTCATCGAA 1327
QY      822 sArgLeuAsnHisLysTyrAsnProValSerLeuProLysAspLeuProAsp---TrpAs 841
Db      1328 GAGGCTGAATCATATAATATAACCCAGTGTACCTTCCCAAGACTTACCCCGACTGGGAGA 1387
QY      841 pTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAlaValLeuCysIleG1 861
Db      1388 TTGGAGACACGGCTGCATCCCTTGCAGAAATATGGAGTTATTGCTGTTCTCTGCATAGA 1447
QY      861 uThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSerAlaTrpLeuPhePh 881
Db      1448 AACAGCCCATATGTTGCTTTTGTGAAGTATGGGAAGGAGATTCCTGGCTGCTCTCTT 1507
QY      881 e-AspSerMetAlaAsp-ArgAspGlyGlyGlnAsnGlyPheAsnIleProGlnValThr 900
Db      1508 TGGACAGCATGGCGATCCGGATGTGGTTCAGAAAGGCTCAACATTCCTCCCAAGTCMCC 1567
QY      901 ProCysProGluValGlyGluTyrLeu-LysMetSer-LeuGluAsp---LeuHisSerL 919
Db      1568 CMTGCCCAAGACTAGGAGAGTACTTGGAAAGATGCTCTCTGGAAAGACCCCTGSAWTVCC 1627
QY      919 euAspSerArgArgIle-GlnGlyCysAlaArgArgLeuLeuCysAspAla-TyrMetCy 938
Db      1628 GGACTCCCAAGGAATCCCAAGGCTGTGCACGAGACTGCTTTGTGATGCCATATATGTG 1687
QY      938 s-MetTyr-GlnSerProThrMetSerLeuTyrLys 949
Db      1688 CCATGTACCAGAGTCCCAACATGAGTTTGTACAAA 1723

```

RESULT 7

```

US-10-264-237-204/c
; Sequence 204, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P4131P1
; CURRENT APPLICATION NUMBER: US/10/264,237
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 204
; LENGTH: 1151
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (15)..(15)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1092)..(1092)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1113)..(1113)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1136)..(1136)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1142)..(1142)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-264-237-204

```

```

Alignment Scores:      2.65e-185      Length:      1151
Pred. No.:      1778.00      Matches:      347
Score:      90.89%      Conservative:      2
Percent Similarity:      90.36%      Mismatches:      15
Best Local Similarity:      35.32%      Indels:      20
Query Match:      17      Gaps:      4
DB:
US-09-671-687A-3 (1-949) x US-10-264-237-204 (1-1151)
QY      534 ProAspSerArgPheAlaSerLeuGlnProValSerAsn---GlnIleGluAtqCysAsn 552
Db      1150 CCTGACCCNAGGTNGCATTTGACGCCCGCTTCANTCCAGATTGGAGCCGCTGTANC 1091
QY      553 SerLeuAla---PheGlyGlyTyrLeuSerGluValValGluGluAsnThrPro---Pro 570
Db      1090 TCTTTAGCCATTGGGAGGGCTTACTTTAAGTGAAGTAGTGGAGGAAATACTCCACCCAA 1031
QY      571 LysMetGluLysGluGlyLeuGluIleMetIle-GlyLysLysLysGlyLleGlnGlyHi 590
Db      1030 AAATGGGAAAAAGAGGCTTTGGAGATAAATGATTGGGGAAGAAAGGCAATCCAGGGTCA 971
QY      590 sTyrAsnSerCysTyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSerValLe 610
Db      970 TTACAATTCTTTTACTTACTTAGACTCAACCTTATTCTGCTTATTTTGCCTTTTAGTCTGTCT 911
QY      610 uAspThrValLeuLeuArgProLysGluLysAsnAspValGluTyrTyrSerGluThrGl 630
Db      910 GGACACTGTGTACTTAGACCCCAAGAAAGACGATGTAGNATATTTATAGTGAACCCA 851
QY      630 nGluLeuLeuArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCysAlaTh 650
Db      850 AGAGCTACTGAGGACAGAAATGTTAATCTCTGAGAAATATATGATATGTGTGTGCCAC 791
QY      650 rLysIleMetLysLeuArgLysIleLeuGluLysValGluAlaAlaSerGlyPheThrSe 670
Db      790 AAAAAATTATGAAACTGAGGAAAAATACTTGAAGAGGTGGAGGCTGCATCAGGATTTACCTC 731
QY      670 rGluGluLysAspProGluGluPheLeuAsnIleLeuPheHisIleLeuArgValGl 690
Db      730 TGAAGAAAAAGATCCTGAGGAATCTTGAATATTTCTGTTTCATCATATTTTAAAGGTAGA 671
QY      690 uProLeuLeuLysIleArgSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIl 710
Db      670 ACCTTTGCTAAAAATAAGATCAGCAGGTCAAAAGGTACAAAGATTGTTACTTCTATCAAT 611
QY      710 ePheMetGluLysAsnGluLysValGlyValProThrIleGlnLeuLeuGluTyrPSe 730
Db      610 TTTTATGGAAAAAATAAGAAAGTTGGCGTTGCCCAATTCAGCAGTTGTTAGAAATGGTC 551
QY      730 rPheIleAsnSerAsnLeuLysPheAlaGluAlaProSerCysLeuIleIleGlnMetPr 750
Db      550 TTTTATCAACAGTAACCTGAAATTTGCAGAGGACCAACATCATGCTGATTTATTCAGATGCC 491
QY      750 oArgPheGlyLysAspPheLysLeuPheLysLysIlePheProSerLeuGluLeuAsnIl 770
Db      490 TCGATTTTGGAAAAAGACTTTAAACTATTTTAAAAAAATTTTCTCTCTCGAAATTTAAATAT 431
QY      770 eThrAspLeuLeuGluAspThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTy 790
Db      430 AACAGATTTTACTTTGAAGACACTCCACAGACAGTCCGATATGTGGAGGGCTTGAATGTA 371
QY      790 rGluCysArgGluCysTyrAspAspProAspIleSerAlaGlyLysIleLysGlnPheCy 810
Db      370 TGAGTGTAGAGAAATGCTTACGAYATCCGGACATCTCAGCTGGAAAAATCAAGAGAGTTTG 311
QY      810 sLysThrCysAsnThrGlnValHisLeuHisProLysArgLeuAsnHisLysTyrAsnPr 830
Db      310 TAAACCTGCACACTCAAGTCCACCTTCATCCGAAGGGCTGATCATATAATATAACCC 251
QY      830 oValSerLeuProLysAspLeuProAspTrpAspTrpArgHisGlyCysIleProCysGl 850
Db      250 AGTGTCACTTTCCCAAGACTTACCCGACTGGAGACTGGAGACACACGCGTGCATCCCTTGCCA 191

```

QY 850 nAsnMetGluLeuPheAlaValIeuCysIleGluThrSerHisTyrValAlaPheValIy 870
 Db 190 GAATATGGAGTATTGCTGCTCTGCATARAACAAGCCACTATGTTGCTTTGTGAA 131
 QY 870 sTyrGlyLysAspSerAlaTriPhePheAspSerMetAlaAspArgAsp----- 888
 Db 130 KTATGGGAAGGACGATTCTGCTGGCYCTCTTTGACAGCATGCCGATCGGATGGTAC 71
 QY 889 -----GlyGlyGln-AsnGlyPheA 895
 Db 70 TGAATAAACGCTTTCTCTGTCATGTTGGCACAAGGGTTCTGCTGGTCAAAAATGNTTCC 11
 QY 895 snIlePro 897
 Db 10 ACATTCCC 3

RESULT 8

US-10-264-237-205
 ; Sequence 205, Application US/10264237
 ; Publication No. US20040009491A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Birse et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PA131P1
 ; CURRENT APPLICATION NUMBER: US/10/264,237
 ; PRIOR FILING DATE: 2002-10-04
 ; PRIOR APPLICATION NUMBER: PCT/US01/16450
 ; PRIOR FILING DATE: 2001-05-18
 ; PRIOR APPLICATION NUMBER: US 60/205,515
 ; PRIOR FILING DATE: 2000-05-19
 ; NUMBER OF SEQ ID NOS: 2876
 ; SOFTWARE: PatentIn Ver. 3.1
 ; SEQ ID NO 205
 ; LENGTH: 1013
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (5)..(5)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (9)..(9)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (307)..(307)
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (948)..(948)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (984)..(984)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1004)..(1004)
 ; OTHER INFORMATION: n equals a,t,g, or c
 US-10-264-237-205

Alignment Scores:

Pred. No.:	4.63e-154	Length:	1013
Score:	1494.00	Matches:	285
Percent Similarity:	98.62%	Conservative:	0
Best Local Similarity:	98.62%	Mismatches:	4
Query Match:	29.68%	Indels:	0
DB:	17	Gaps:	0

US-09-671-687a-3 (1-949) x US-10-264-237-205 (1-1013)

QY 358 SerGlnProGlnSerLysSerLysAsnThrTyrIleAspGluValAlaGluAspPro 377
 Db 12 TCACACCAACCAATCCAAATCAAAAATACATGGTACATTGATGAAGTTGCAGAGACCT 71
 QY 378 AlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProProLeuGlnPro 397
 Db 72 GCAAAATCTCTTACAGAGATATCTACAGACTTTTGACCGTCTTCCACCACCACTCCAGCT 131
 QY 398 ProProValAsnSerLeuThrGluAsnArgPheHisSerLeuProPheSerLeuThr 417
 Db 132 CTTCTCTGTGAACCTACCTGACCACCGAGAACAGATTCCACTCTTTACCACTTCAGTCTCACC 191
 QY 418 LysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGlnSer 437
 Db 192 AGATGCCCAATACCAATGGAAGTATTGGCCACAGTCCACTTTCTGTGAGCCAGTCT 251
 QY 438 ValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetProPro 457
 Db 252 GTAATGGAAGAGCTAAACACTGCACCGTCCAAGAGAGTCCACCCTTTGGCCATGCNTCT 311
 QY 458 GlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPhe 477
 Db 312 GGGAACTCATCGTCTAGAAGTGGGCTCATTTGGCTGAAGTTAAGGAGAACCTCTCTTC 371
 QY 478 TyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGlyLeu 497
 Db 372 TATGGGGTAAATCCGTTGGATCGGTGAGCCACCAAGGACTGAATGAAGTCTCGCTGGACTG 431
 QY 498 GluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyrPhe 517
 Db 432 GAACTGGAAGATGAGTGTGAGCTGTACGGATGGAACCTTCAGAGGACCTCGGTATTTC 491
 QY 518 ThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSerArg 537
 Db 492 ACCTGTGCCCTGAAGAARGCGCTGTTGTGAACCTGAAGAGCTGCAGGCTGACTCTAGG 551
 QY 538 PheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPheGly 557
 Db 552 TTTGTCATCATTTGAGCGCGTTTCCAATCAGATTGAGCGCTGTAACTCTTTAGCATTTTGA 611
 QY 558 GlyTyrLeuSerGluValValGluAsnThrProProLysMetGluLysGluGlyLeu 577
 Db 612 GGCTACTTAAGTGAAGTAGTAGAAGAAATACTCCACCAAAATGGAAAAGAGGCTTG 671
 QY 578 GluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeuAsp 597
 Db 672 GAGATAATGATTGGGAAGAAGAGCATCCAGGTCATTACAATCTCTGTACTTAGAC 731
 QY 598 SerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArgPro 617
 Db 732 TCAACCTTATTCTGCTTATTCTTTTAGTTCTGTTCTRGACACTGTGTTACTTAGACCC 791
 QY 618 LysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGluIle 637
 Db 792 AAAGAAAAGAACGATGAGAATATTATAGTAAACCAAGAGCTACTGAGGACAGAAAT 851
 QY 638 ValAsnProLeuArgIleTyrGlyTyr 646
 Db 852 GTTAATCTCTGTGAAATTAATKATAC 878

RESULT 9

US-09-969-034-4253/c
 ; Sequence 4253, Application US/09969034
 ; Publication No. US20040110668A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Burgess, Christopher C.
 ; APPLICANT: Aetle, Jon H.
 ; APPLICANT: Carroll, Eddie III
 ; APPLICANT: Catino, Theodore J.
 ; APPLICANT: Dwivedi, Poornima
 ; APPLICANT: Molino, Gary A.
 ; APPLICANT: Thiagalingam, Arunthathi
 ; APPLICANT: Lewis, Marcia E.

RESULT 10
; TITLE OF INVENTION: Nucleic Acid Sequences Differentially
; FILE REFERENCE: 1657/1032
; CURRENT APPLICATION NUMBER: US/09/969,034
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/237,271
; NUMBER OF SEQ ID NOS: 4494
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4253
; LENGTH: 617
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 517, 586
; OTHER INFORMATION: n = A,T,C or G
US-09-969-034-4253

Alignment Scores:
Pred. No.: 3.67e-111 Length: 617
Score: 1102.00 Matches: 203
Percent Similarity: 98.54% Conservative: 0
Best Local Similarity: 98.54% Mismatches: 2
Query Match: 21.89% Indels: 2
DB: 11 Gaps: 0

US-09-671-687A-3 (1-949) x US-09-969-034-4253 (1-617)

QY	703	GlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGluLysValGlyValProThr	722
Db	616	CAAGATTGTTATTCTATCAAAATTTTATGNA-AAAAATGAGAAAGTTGGCGTTCACCA	558
QY	723	IleGlnLeuLeuGluThrPheSer-PheIleAsnSerAsnLeuLysPheAlaGluAlaPr	742
Db	557	ATTCACAGTGTAGAAATGCTTTTATCAACAGTAACNTGAAATTTGCAGAGGCACC	498
QY	742	oSerCysLeuIleIleGlnMetProArgPheGlyLysAspPheLysLeuPheLysLysI	762
Db	497	ATCATGCTGATTATTTCAGATGCTCGATTGGAAAAGACTTTTAACTATTAAAAAAT	438
QY	762	ePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAspThrProArgGlnCysAr	782
Db	437	TTTTCTCTCTCGAAATTAATAATAACAGATTTTACTTGAAGACACTCCACAGACAGTCCG	378
QY	782	gileCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyrAspAspProaspIleSe	802
Db	377	GATATGTGGAGGCTTGCAATGTATGAGTGTAGAGATGCTACGACGATCCCGACATCTC	318
QY	802	xAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGlnValHisLeuHisProLy	822
Db	317	AGCTGGAATAATCAACAGCTTTTGTAAACCTGCAACACTCAAGTCCACCTTCATCCGAA	258
QY	822	sArgLeuAsnHisLysTyrAsnProValSerLeuProLysAspLeuProAspTrpAspTr	842
Db	257	GAGCTGGAATCAATAATAACCCAGTGTACCTTCCAAAGACTTACCCGACTGGGACTG	198
QY	842	pArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAlaValLeuCysIleGluTh	862
Db	197	GAGACACGGCTGATCCCTTGCAGAAATATGGAGTTATTTCCTGCTCTGCATAGAAAC	138
QY	862	rSerHisTyrValAlaPheValLysTyrGlyLysAspAspSerAlaTrpLeuPhePheSe	882
Db	137	AGCCACTATGTTGCTTTTGTGAAGTATGGGAAGGACGATTTCGCCCTGCTCTCTTTGA	78
QY	882	pSerMetAlaAspArgAspGlyGlnAsnGlyPheAsnIleProGlnValThrProCy	902
Db	77	CAGCATGGCCGATCGGATGCTGTGAGATGCTTCAACATTCCTCAAGTCAACCCCATG	18
QY	902	sProGluValGlyGlu	907
Db	17	CCGAGAAGTAGGAGAG	2

US-09-969-034-4254/c
; Sequence 4254, Application US/09969034
; Publication No. US20040110668A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Astle, Jon H.
; APPLICANT: Carroll, Eddie III
; APPLICANT: Catino, Theodore J.
; APPLICANT: Dwivedi, Poornima
; APPLICANT: Molino, Gary A.
; APPLICANT: Thiagalingam, Arunthathi
; APPLICANT: Lewis, Marcia E.
; TITLE OF INVENTION: Nucleic Acid Sequences Differentially
; FILE REFERENCE: 1657/1032
; CURRENT APPLICATION NUMBER: US/09/969,034
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/237,271
; NUMBER OF SEQ ID NOS: 4494
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4254
; LENGTH: 557
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 429, 468, 523
; OTHER INFORMATION: n = A,T,C or G
US-09-969-034-4254

Alignment Scores:

Pred. No.: 1.24e-86 Length: 557
Score: 879.00 Matches: 170
Percent Similarity: 91.44% Conservative: 1
Best Local Similarity: 90.91% Mismatches: 12
Query Match: 17.46% Indels: 5
DB: 11 Gaps: 2

US-09-671-687A-3 (1-949) x US-09-969-034-4254 (1-557)

QY	725	GlnLeuLeuGluThrPheSerPheIleAsnSer---AsnLeuLysPheAlaGluAlaPro---	742
Db	557	CAGTGTAGATGGTGTATTAATCAACAGTTAACTGAAATTTTGCAAAAGGCACCAT	498
QY	743	SeCysLeuIleIleGlnMetProArgPheGlyLysAspPheLysLeuPheLysLysIle	762
Db	497	CAATGTTGATTATTCAGATGCTCGATT-NGAAAAGACTTTTAACTATTAAAAAAT	439
QY	763	PheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAspThrProArgGlnCysArg	782
Db	438	TTTCTCTTTTNGAATTAATAATAACAGATTTTACTTGAAGACACTCCACAGACAGTCCGG	379
QY	783	IleCysGlyGlyLeu-AlaMetTyrGluCysArgGluCysTyrAsp-AspProaspIleSe	802
Db	378	ATATGTGGAGGCTTTGCAATGTATGAGTGTAGAGATGTTACGACGATCCGACATCT	319
QY	802	erAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGlnValHisLeuHisProL	822
Db	318	CAGCTGGAATAATCAACAGCTTTTGTAAACCTTCAACACTCAAGTCCACCTTCATCCGA	259
QY	822	YsArgLeuAsnHisLysTyrAsnProValSerLeuProLysAspLeuProAspTrpAspTr	842
Db	258	AGAGGCTGAATCAATAATAATAACCCAGTGTCTTCCAAAGACTTACCCGACTGGGACT	199
QY	842	rpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAlaValLeuCysIleGluT	862
Db	198	GGAGACACGGCTGCATCCCTTGCAGAAATATGAGTTATTTCCTGCTCTGCATAGAA	139
QY	862	hrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSerAlaTrpLeuPhePheA	882
Db	138	CAAGCCACTATGTTGCTTTTGTGAAGTATGGGAAGGACGATTTCCTGCTCTCTCTTTG	79

```
QY      882  spSerMetAlaaspArgaspGlyGlyGlnAsnGlyPheAsnIleProGlnValThrProC 902
Db      78  ACAGCATGCCGATCGGATGGTGGTGAATGCTTCAACATTCCTCAAGTCACCCCAT 19
QY      902  yaProGluValGlyGlu 907
Db      18  GCCCAGAGTAGGAGAG 2

RESULT 11
US-09-983-965-1633
; Sequence 1633, Application US/09983965
; Patent No. US20020137160A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 37-21(10297)C
; CURRENT APPLICATION NUMBER: US/09/983,965
; PRIOR FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: US 09/465,231
; PRIOR FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: US 60/113,678
; PRIOR FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 5912
; SEQ ID NO 1633
; LENGTH: 376
; TYPE: DNA
; ORGANISM: Bos taurus
; FEATURE:
; OTHER INFORMATION: Clone ID: 47-LIB2809-020-Q1-E1-D4
US-09-983-965-1633

Alignment Scores:
Pred. No.:      5,38e-59      Length:      376
Score:          626.00      Matches:    117
Percent Similarity: 94.35%      Conservative: 0
Best Local Similarity: 94.35%      Mismatches: 7
Query Match:      12.44%      Indels:     0
DB:              9          Gaps:     0

US-09-671-687A-3 (1-949) x US-09-983-965-1633 (1-376)

QY      703  GlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGluLysValGlyValProThr 722
Db      5   CAAGATAGTTACATCTATCAAAATTTTATGGAAGAAAATGAGAAAGTTGGAGTTCCTACA 64
QY      723  IleGlnGlnLeuLeuGluTrpSerPheIleAsnSerAsnLeuLysPheAlaGluAlaPro 742
Db      65  ATTACAGCAATTATTAGAATGTTCTTTTATCAACAGTAACCTGAATTTGCTGAGGCACCA 124
QY      743  SerCysLeuIleIleGlnMetProArgPheGlyLysAspPheLysLeuPheLysLysIle 762
Db      125  TCATGCTCTGATTATTACAGATGCTCGATTTTGGAAAAGACCTTTAAACTCTTTAAAAAATT 184
QY      763  PheProSerLeuLeuLeuAsnIleThrAspLeuLeuGluAspThrProArgGlnCysArg 782
Db      185  TTTCTCTCTCTGGAACATAATAACAGATTATCTTGAAGACACCCCGGAGTGCCTCA 244
QY      783  IleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyrAspAspProAspIleSer 802
Db      245  ATATGTTGGAGGGCTTGCATATGATGAGTGATGAGATGCTATGATGACCTTGACATCTCA 304
QY      803  AlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGlnValHisProLys 822
Db      305  GCTGGCACAATCAAGCAGATTGTTGTAACACCTGCAATGCTCAAGTCCACCTTCAGTCTAAG 364
QY      823  ArgLeuAsnHis 826
Db      365  AGGCTGAATCAT 376
```

```
RESULT 12
US-09-864-761-18145/c
; Sequence 18145, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 18145
; LENGTH: 309
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007728.1
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.98
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.92
; OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.89
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
; OTHER INFORMATION: EST HUMAN HIT: A1130924.1, EVALUATE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: Q10427, EVALUATE 1.70e-01
; OTHER INFORMATION: NT HIT: G111430965, EVALUATE 0.00e+00
US-09-864-761-18145
```

Alignment Scores:
Pred. No.: 3,38e-50 Length: 309
Score: 545.00 Matches: 102
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.83% Indels: 0
DB: 0 Gaps: 0

US-09-671-687A-3 (1-949) x US-09-864-761-18145 (1-309)

QY 387 AspPheAspArgSerProProLeuGlnProProValAsnSerLeuThrThrGlu 406
DB 308 GACTTGGACGGTCTTCCACACACCTCCAGCCCTCTCTGTGAACTCATGACCACCGAG 249

QY 407 AsnArgPheHisSerLeuProPheSerLeuThrLysMetProAsnThrAsnGlySerIle 426
DB 248 AACAGATTCCACTTTTACCAATTCAGTCTCACCACAGATGCCAATACCAATGGAAGTAAT 189

QY 427 GlyHisSerProLeuSerLeuSerAlaGlnSerValMetGluLulLeuAsnThrAlaPro 446
DB 188 GGCCACAGTCCACTTTCTGTGTGAGCCAGTCTGTAAATGGAAGAGCTAAACACTGCACCC 129

QY 447 ValGlnGluSerProProLeuAlaMetProProGlyAsnSerHisGlyLeuGluValGly 466
DB 128 GTCCACAGAGTCCACCTTGGCCATGCCCTCTCTGGGAACCTCATGTCTAGAAGTGGGC 69

QY 467 SerLeuAlaGluValLysGluAsnProProPheTyrGlyValIleArgTrpIleGlyGln 486
DB 68 TCATTGGCTGAAGTTAAGGAGAACCTCTCTTCTATGGGGTAAATCGTTGGATCGTGCAG 9

QY 487 ProPro 488
DB 8 CCACCA 3

RESULT 13
US-09-864-761-1386/c
; Sequence 1386, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeonica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 1386
; LENGTH: 483
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007728.1
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.98
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.92
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.89
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4
US-09-864-761-1386

Alignment Scores:
Pred. No.: 3,05e-46 Length: 483
Score: 512.00 Matches: 98
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.17% Indels: 0
DB: 9 Gaps: 0

US-09-671-687A-3 (1-949) x US-09-864-761-1386 (1-483)

QY 402 SerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeuThrLysMetProAsn 421
DB 483 TCCTCACCACCCAGAACAGATTCCACTCTTTACCATTCAGTCTCACCAGATGCCCAAT 424

QY 422 ThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGlnSerValMetGluGlu 441
DB 423 ACCAATGGAAGTATTGGCCACAGTCCACTTCTCTGTGACCCAGTCTGTAAATGGAAGAG 364

QY 442 LeuAsnThrAlaProValGlnGluSerProProLeuAlaMetProProGlyAsnSerHis 461
DB 363 CTAACACTGCACCCCGTCCAAAGAGAGTCCACCCTTGGCCATGCCCTCTGGGAACCTCAT 304

QY 462 GlyLeuGluValGlySerLeuAlaGluValLysGluAsnProProPheTyrGlyValIle 481
DB 303 GGCTAGAGTGGGCTCATTTGGCTGAAGTTAAGGAGAACCTCTCTTCTATGGGGTAAATC 244

QY 482 ArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGlyLeuGluLeu 499
DB 243 CGTTGGATCGTCCAGCCAGGACTGAATGAAGTGTCTGCTGACTGGAACTG 190

RESULT 14
US-09-918-995-35878
; Sequence 35878, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20

; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 35878
; LENGTH: 425
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(425)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-35878

Alignment Scores:
Pred. No.: 2,67e-35 Length: 425
Score: 412.00 Matches: 89
Percent Similarity: 63.12% Conservative: 0
Best Local Similarity: 63.12% Mismatches: 1
Query Match: 8.18% Indels: 52
DB: 10 Gaps: 1

US-09-671-687A-3 (1-949) x US-09-918-995-35878 (1-425)

QY 565 GluGluAsnThrProProLysMetGluLysGluGlyLeuGluMetIleGlyLysLys 584
Db 3 GAAGAAATACTCCACCAAAATGGAAGAGGCTTGGAGATATGATGGGAAGAG 62
QY 585 LysGlyIleGlnGlyHisTyrAsnSerCysTyrLeuAspSerThrLeuPheCysLeuPhe 604
Db 63 AAAGGCATCCAGGGTCATTACAAATCTTGTTACTTAGACTCAACCTTAATCTGCTTATT 122
QY 605 AlaPheSerSerValLeuAspThrValLeuLeuArgProLysGluLysAsnAspValGlu 624
Db 123 GCTTTTAGTCTGTTCTTGACACTGTGTTACTTAGACCCCAAGAAAGAACGATGTAGAA 182
QY 625 TyrTyrSerGluThrGlnGluLeuLeuArgThrGluIleValAsnProLeuArgIle--- 643
Db 183 TATTATAGTGAACCCACAGAGCTACTGAGACAGAAATTTGTAATCTCTGAGAAAT-TAA 241
QY 643 ----- 643
Db 242 TTGATACTGGAACAAGACAGGTCAACCCGAACTCTAAGGCATGTTAACACCTCCTTCC 301
QY 643 ----- 643
Db 302 TAGAAGAGAGGAAGATAAACATCTCTTGGATCCTATTGTTAGCCAGTTAAGAAATCATA 361
QY 644 -----TyrGlyTyrValCysAlaThrLysIleMet 653
Db 362 GTTAATCACACTTTCTTACCCCTGTCCAGATATGGATATGTGTGCCCAAAAAATTATG 421
QY 654 Lys 654
Db 422 AAA 424

RESULT 15

US-10-152-319A-1764
; Sequence 1764, Application US/10152319A
; Publication No. US20040072160A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Higgs, Brandon
; APPLICANT: Castle, Arthur
; APPLICANT: Elashoff, Michael
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5089-US
; CURRENT APPLICATION NUMBER: US/10/152,319A
; CURRENT FILING DATE: 2002-05-22
; PRIOR APPLICATION NUMBER: US 60/292,335
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/297,523
; PRIOR FILING DATE: 2001-06-13

; PRIOR APPLICATION NUMBER: US 60/298,925
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,810
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/303,807
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/303,808
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/315,047
; PRIOR FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: US 60/324,928
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/330,867
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: US 60/330,462
; PRIOR FILING DATE: 2001-10-22
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2221
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1764
; LENGTH: 4847
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. NM_021997
US-10-152-319A-1764

Alignment Scores:
Pred. No.: 7,68e-06 Length: 4847
Score: 160.50 Matches: 93
Percent Similarity: 32.76% Conservative: 40
Best Local Similarity: 22.91% Mismatches: 143
Query Match: 3.19% Indels: 130
DB: 17 Gaps: 16

US-09-671-687A-3 (1-949) x US-10-152-319A-1764 (1-4847)

QY 301 ProGluSerValThrGlnGluArg-----ArgProProLysLeuAlaPheMet 316
Db 272 CCTAGTGATCATCCATGACAGAGCCAGTGGCGCTGAAGCCCTC----- 316
QY 317 SerArgGlyValGlyAspLysGlySerSerSerHisAsnLysProLysAlaThrGlySer 336
Db 317 -----GGCCGAGGGGAGACACTCCAGTCCCGTGGGTAGGCCCTCCATC- 361
QY 337 ThrSerAspProGlyAsnArgArgSerGluLeuPheTyrThrLeuAsnGlySerSerVal 356
Db 362 -----GGGTCTGCTCATCTTCTGTGTGCATCGGCGAGCGGCTCCAAAGAA 409
QY 357 AspSerGln-----ProGlnSerLysSerLysAsnThrTrpTyr 369
Db 410 GGGTCTCTCCCTGCACAAGCAGGCGTGGGCGCTCTCTCGGTGGAGCCCACT- 463
QY 370 IleAspGluValAlaGluAspProAlaLysSerLeuThrGluLysSerThrAspPhe--- 388
Db 464 -----ACTGTCTCCGAGAAGCCAGGTCCAAAGCAGCCGAGTGGGTGATGACTTCTTG 517
QY 389 -----AspArgSerSerProProLeuGlnProProValAsnSer 402
Db 518 GGAGACTTCGTGTGGCGAGAGGGTGTGGGTGAATGCGCTGAAGCCCGCGGTGGTCAG 577
QY 403 LeuThrThrGluAsnArgPhe----- 409
Db 578 TACCTGGGTGAGACACACAGTTTGGCCAGGCCAGTGGCGCGGTGGTCTTGGATGATCCT 637
QY 409 ----- 409
Db 638 GTGGCAAGAACCATGGAGCTGTGGGTGGGTGGCTGCTTCCAGTCCAGCTCTACAG 697
QY 410 -----HisSerLeuProPheSerLeuThrLysMetPro-----AsnThrAsnGlySer 425
Db 698 GGCATCTTCAACAGCGCCCTCCAAAGCTGACCCGCGCAGCCGCGAGCTGAGGGGCTCCGCGCAGT 757

```
QY 426 IleGlyHisSerProLeuSerAlaGlnSerValMetGluLeuAsnThrAla 445
Db 758 GACGGCCACTCAGTGAATCCCTCACCGCCCGAGACCTGCTCCTCATTTCTGGCACGGCC 817
QY 446 ProValGlnGluSerProLeu----- 453
Db 818 -----ACGCCCCCACTCACCGCCCGAGTATCCCCCTGGGGAGAGTGTTC 865
QY 454 -----AlaMetProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluVal 471
Db 866 AACAGCTCCGTGAAGACGGCAATGAGTCTGGCTCCAACTCTCGGACAGTGTCTGTG 925
QY 472 LysGluAsnProProPheTyr----- 478
Db 926 AAGCGTGGTGACAGGACCTCCACCTGGGAGACCGTGTCTGTTGTTGGGACGAGACC 985
QY 479 GlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGlyLeuGlu 498
Db 986 GGTGTGTTAGCATATGTGGGAGACGGACTTTGCCAAGGCCAGTGGTGTGTGGAG 1045
QY 499 LeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyrPheThr 518
Db 1046 CTG---GACGAGCCCTTGGGAGACGATGGGCGAGTGGCAGCGCACCGACTTCCAG 1102
QY 519 CysAlaLeuLysLysAlaLeuPheValLysLeu----- 529
Db 1103 TGCCCAACCAAGTTCGTCTCTTTGACCAATCCACAAGTTCATCCGAATTGGCTTCCCA 1162
QY 530 -----LysSerCysArgProAspSerArgPheAla----- 539
Db 1163 TTTACCACTCCAGCCCAAGGCCAAGACACCAAGCGCATGGCTATGGGTGTCTCAGCCTTG 1222
QY 540 SerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAla-----PheGly 557
Db 1223 ACCCAGAGCCCGAGCAGTCTTCCATCAGCTCTGTGAGCTCTGTGGCTCTCTCTGTGT 1282
QY 558 GlyTyrLeuSerGlu-----ValValGluGluAsnThrProProLysMetGluLys 574
Db 1283 GGCGGCGGAGCGTAGTGGCCTGCTCACAGAGACCTTTTCACGCTATGCCCGGAGATC 1342
QY 575 GluGlyLeuGluIleMetIleGlyLysLysLysGlyIleGlnGlyHisTyr-AsnSerCy 594
Db 1343 TCAGGCACCCACAGCCCTTGCAGGAGGCACCTGAAGGAGAAACAGCAGCAGCATCGAACAGCTG 1402
QY 594 sTyrLeuAspSerThr 599
Db 1403 CTGGCTGAGCGTGA 1418
```

Search completed: April 18, 2005, 22:42:32
Job time : 1199 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 15:16:44 ; Search time 65 Seconds
(without alignments)
7476.356 Million cell updates/sec

Title: US-09-671-687A-3

Perfect score: 5034

Sequence: 1 MSSGLWSQEKVTSPTWEERI.....RLLCDAYMCVQSPMTSLYK 949

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4971.5	98.8	956	1 CYLD HUMAN	Q9nqc7 homo sapien
2	4718.5	93.7	952	1 CYLD MOUSE	Q80tq2 mus musculus
3	4712	93.6	953	2 Q6H62	Q6h62 rattus norv
4	1207	24.0	778	2 Q70414	Q7q414 anopheles g
5	1136.5	22.6	551	2 Q8IPC3	Q8ipc3 drosophila
6	1136.5	22.6	639	2 Q8IPC5	Q8ipc5 drosophila
7	1134	22.5	550	2 Q8IPC4	Q8ipc4 drosophila
8	1132.5	22.5	639	2 Q8SYF0	Q8syf0 drosophila
9	1092.5	21.7	454	2 Q6TXJ6	Q6txj6 rattus norv
10	856	17.0	517	2 Q9VL04	Q9vl04 drosophila
11	619	12.3	1144	2 Q7JMS4	Q7jms4 caenorhabdi
12	183.5	3.6	258	2 Q7PUA3	Q7pja3 anopheles g
13	153.5	3.0	528	2 Q6BU87	Q6bu87 debaryomyce
14	151	3.0	653	2 Q6P5Z9	Q6p5z9 homo sapien
15	151	3.0	1427	1 REST HUMAN	P30622 homo sapien
16	149.5	3.0	1046	2 Q5S156	Q5s156 rattus norv
17	149	3.0	1046	2 Q9Z0H8	Q9z0h8 mus musculus
18	147.5	2.9	449	2 Q86WU4	Q86wu4 homo sapien
19	146	2.9	350	2 Q6PIA3	Q6pia3 homo sapien
20	145	2.9	547	2 Q96DZ5	Q96dz5 homo sapien
21	145	2.9	1556	2 Q9VRP1	Q9vrp1 drosophila
22	143.5	2.9	495	2 Q6FM69	Q6fm69 candida gla
23	143.5	2.9	1047	2 Q9EP81	Q9ep81 mus musculus
24	143	2.8	547	2 Q8MWL1	Q8mw11 homo sapien
25	142.5	2.8	1606	2 Q7PVQ8	Q7pvq8 anopheles g
26	142	2.8	1012	2 Q7T5I9	Q7tsi9 mus musculus
27	141	2.8	1012	2 Q8CHU1	Q8chu1 mus musculus
28	141	2.8	1429	2 Q6DFB6	Q6dfb6 xenopus lae
29	140.5	2.8	429	2 Q96C99	Q96c99 homo sapien
30	140	2.8	924	2 Q7VDY2	Q7vdy2 prochloroco
31	139.5	2.8	1921	2 Q86BR0	Q86br0 drosophila

ALIGNMENTS

RESULT 1

ID	CYLD_HUMAN	STANDARD;	PRT;	956 AA.
AC	Q9NQC7; Q94934; Q7L3N6; Q96EH0; Q9NZX9;			
DT	25-OCT-2004 (Rel. 45, Created)			
DT	25-OCT-2004 (Rel. 45, Last sequence update)			
DT	25-OCT-2004 (Rel. 45, Last annotation update)			
DE	Probable ubiquitin carboxyl-terminal hydrolase CYLD (EC 3.1.2.15)			
DE	(Ubiquitin thiolesterase CYLD) (Ubiquitin-specific processing protease			
DE	CYLD) (Deubiquitinating enzyme CYLD) (HSPC057).			
GN	Name=CYLD; Synonyms=CYLD1, KIAA0849;			
OS	Homo sapiens (human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORM 1), DISEASE, AND TISSUE SPECIFICITY.			
RX	MEDLINE=20296617; PubMed=10835629; DOI=10.1038/76006;			
RA	Bignell G., Brown C., Biggs P.J., Lakhani S.R., Jones C., Hansen J.,			
RA	Blair E., Hofmann B., Siebert R., Turner G., Evans D.G.,			
RA	Schander-Stumpel C., Beemer F.A., Van Den Ouweland A., Halley D.,			
RA	Delpech B., Cleveland M.G., Leigh I., Leisti J., Rasmussen S.,			
RA	Wallace M.R., Fenske C., Banerjee P., Oiso N., Chaggar R., Merrett S.,			
RA	Leonard N., Huber M., Hohl D., Chapman P., Buern J., Swift S.,			
RA	Smith A., Ashworth A., Stratton M.R.;			
RT	"Identification of the familial cylindromatosis tumor suppressor			
RT	gene.";			
RL	Nat. Genet. 25:160-165(2000).			
RN	[2]			
RP	SEQUENCE FROM N.A. (ISOFORM 2).			
RC	TISSUE=Brain;			
RC	MEDLINE=99156230; PubMed=10048485;			
RX	Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hiroseawa M.,			
RA	Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;			
RT	"Prediction of the coding sequences of unidentified human genes. XII.			
RT	The complete sequences of 100 new cDNA clones from brain which code			
RT	for large proteins in vitro.";			
RL	DNA Res. 5:355-364(1998).			
RN	[3]			
RP	SEQUENCE FROM N.A. (ISOFORM 2).			
RC	TISSUE=Uterus;			
RX	MEDLINE=22398257; PubMed=12477932; DOI=10.1073/pnas.242603899;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.L., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RA	Rana S.S., Loquellano N.A., Peters G.J., Carninci P., Prange C.,			
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,			

32	139	2.8	446	2	Q757P1	Q757p1 ashbya goss
33	138.5	2.8	1172	2	Q7RFU4	Q7rfu4 plasmodium
34	138.5	2.8	1937	2	Q6A062	Q6a062 mus musculus
35	138	2.7	724	2	Q8BWD1	Q8bwd1 mus musculus
36	138	2.7	1024	2	Q14527	Q14527 homo sapien
37	138	2.7	1391	2	Q922J3	Q922j3 mus musculus
38	137.5	2.7	992	2	Q6CIK8	Q6cik8 varrowia li
39	137	2.7	868	2	Q9SDN6	Q9sdn6 nicotiana t
40	136.5	2.7	359	2	Q7PVQ6	Q7pvq6 anopheles g
41	136	2.7	887	2	Q6FQF0	Q6fqf0 candida gla
42	135.5	2.7	2273	2	Q31152	Q31152 neisseria m
43	135.5	2.7	2273	2	Q7DD08	Q7ddd8 neisseria m
44	135	2.7	1046	2	Q9UDT6	Q9udt6 homo sapien
45	134.5	2.7	657	2	Q7Z3N8	Q7z3n8 homo sapien

Db 121 KGLQIDVGPVKVQLSGEKEPGVVRFRGPLLAERTVSGIFPGVELLEGRGQGTG 180
 QY 180 YGKQLFQCDGCG-FVALDKLELIEDDTALESYAGPGDTWQVLPPELINSRVS LKG 238
 Db 181 YGKQLFQCDGCGFVALDKLELIEDDTALESYAGPGDTWQVLPPELINSRVS LKV 240
 QY 239 GTTIESGTIVFCVLPKGSLSGYFVGVDMNPFGNWDGRFDGV-LCSFACVESTILLHIN 297
 Db 241 GETIESGTIVFCVLPKGSLSGYFVGVDMNPFGNWDGRFDGVOLCSFACVESTILLHIN 300
 QY 298 DIIP---ESVTOERRPPKLAFLMRSGVGDGKSSSHNPKPATGSTSDPGNR-RSLSFYTLNG 353
 Db 301 DIIPALSESVTOERRPPKLAFLMRSGVGDGKSSSHNPKPATGSTSDPGNRNRSLSFYTLNG 360
 QY 354 SSVDSQPSQSKNTWYIDEVAEDPAKSLTEISTDFDRSSPPLOPPVNSLTITNRRFHS L P 413
 Db 361 SSVDSQPSQSKNTWYIDEVAEDPAKSLTEISTDFDRSSPPLOPPVNSLTITNRRFHS L P 420
 QY 414 FSLTKMPNTNGIGHSPLSLSAQSVMEELNTAPVQSPPLAMPNGSHGLEVGS LAE VKE 473
 Db 421 FSLTKMPNTNGIGHSPLSLSAQSVMEELNTAPVQSPPLAMPNGSHGLEVGS LAE VKE 480
 QY 474 NPPFYGVIRWIGOPPLGNEVLAGLEDEDCAGCTDGTFRGTRYFTCALKKALFVKLS CR 533
 Db 481 NPPFYGVIRWIGOPPLGNEVLAGLEDEDCAGCTDGTFRGTRYFTCALKKALFVKLS CR 540
 QY 534 PDSRFASLPVSNQIRCNLSLAFGGVYLSEVVENTPPKMEKEGLETWI GKKGIQGHYNS 593
 Db 541 PDSRFASLPVSNQIRCNLSLAFGGVYLSEVVENTPPKMEKEGLETWI GKKGIQGHYNS 600
 QY 594 CYLDTSLFCLFAPSSVLDTVLLRPKEKNDVEYVYSETQELLRTBI VNP LRIYGVVCA TKIM 653
 Db 601 CYLDTSLFCLFAPSSVLDTVLLRPKEKNDVEYVYSETQELLRTBI VNP LRIYGVVCA TKIM 660
 QY 654 KLRKILEKVEAASGFTSEKDEEFNLILFHHILRVEPLILKRSAGQVDCYFYQIFME 713
 Db 661 KLRKILEKVEAASGFTSEKDEEFNLILFHHILRVEPLILKRSAGQVDCYFYQIFME 720
 QY 714 KNEKGVPTIQOLLEWSFNSLNKFAEAPSLIIQMPRFQKPKFKIPPSLELNI TDL 773
 Db 721 KNEKGVPTIQOLLEWSFNSLNKFAEAPSLIIQMPRFQKPKFKIPPSLELNI TDL 780
 QY 774 LEDTPRQCRICGLAMCYECREYDDPDISAGKIKQFCCTKTQVHLHPKRLNHNKYNPVSL 833
 Db 781 LEDTPRQCRICGLAMCYECREYDDPDISAGKIKQFCCTKTQVHLHPKRLNHNKYNPVSL 840
 QY 834 PKDLPDWDRHGCIPQNNMELFAVLCTIETSHYVAFVKYKDDSAWLPFDSMADRDGGQNG 893
 Db 841 PKDLPDWDRHGCIPQNNMELFAVLCTIETSHYVAFVKYKDDSAWLPFDSMADRDGGQNG 900
 QY 894 FNIPOVTPCPEVGEYKMSLEDLHSLDSRRIOGCARRLLCDAYMCMVQSP TMSLYK 949
 Db 901 FNIPOVTPCPEVGEYKMSLEDLHSLDSRRIOGCARRLLCDAYMCMVQSP TMSLYK 956

RESULT 2
 CYLD MOUSE
 ID CYLD MOUSE STANDARD; PRT; 952 AA.
 AC Q80TQ2; Q80VB3; Q8BX23; Q8YBL9; Q8CGB0;
 DT 25-OCT-2004 (Rel. 45, Created)
 DT 25-OCT-2004 (Rel. 45, Last sequence update)
 DE Probable ubiquitin carboxyl-terminal hydrolase CYLD (EC 3.1.2.15)
 DE (Ubiquitin thiolesterase CYLD) (Ubiquitin-specific processing protease CYLD) (Deubiquitinating enzyme CYLD).
 GN Name=Cyld; Synonyms=Cyld1, Kiaa0849;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 2).

RC TISSUE=Brain;
 RA MEDLINE=22579291; PubMed=12693553;
 RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Aizawa H., Yuasa S.,
 RA Nakajima D., Nagase T., Ohara O., Koga H.;
 RT "Prediction of the coding sequences of mouse homologues of KIAA gene:
 RT II. The complete nucleotide sequences of 400 mouse KIAA-homologous
 RT cDNAs identified by screening of terminal sequences of cDNA clones
 RT randomly sampled from size-fractionated libraries.";
 RL DNA Res. 10:35-48(2003).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 3), AND SEQUENCE OF 1-620 FROM N.A.
 RP (ISOFORM 1).
 RC STRAIN=C57BL/6J; TISSUE=Cerebellum, and Hypothalamus;
 RA MEDLINE=22334683; PubMed=12466851; DOI=10.1038/nature01266;
 RA Okazaki I., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nikaido I., Osato Y., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Bradt D., Brueic V., Chothia C., Corbani L.E., Cousins S.,
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Konegaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wyshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hahizume W., Imotani K., Ishii Y., Itoh M., Kigawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC STRAIN=FVB/N; TISSUE=Mammary gland;
 RA MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
 RA Schnurch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -I- FUNCTION: Negative regulator of TRAF2 and NF-kappa-B signaling
 CC pathway. Has deubiquitinating activity that is directed towards
 CC non-lys-48-linked polyubiquitin chains. The inhibition of NF-
 CC kappa-B activation is mediated at least in part, by the
 CC deubiquitination and inactivation of TRAF2 and, to a lesser
 CC extent, TRAF6 (By similarity).


```
Db 94 -----SSDKKTQTPP-----SSMALGNGNGTAGVAASGGSTSSNS 129
QY 340 PGNRRSELFYTLNGSSVDSQ-POSKSKNTWYIDVDEADPAKSIETSTDFDRSSPPLQPP 398
Db 130 TANSSGSHYSHNGSSASGRYERNKSLNPLAHGMAKQOQETVAGVRGGRAG-DAKPA 188
QY 399 PYN--SLTTENRPHSL-----PFLTKMPNTNG---SIGHPSLSAQSVMEELNTP 446
Db 189 TVDVATMTKLNDAATANGISIGFADGTSHTPTANGLVPRIG-----SISPESDSGGPAAS 244
QY 447 VQESPLAMPNGSHGLEVCSLAEV-----KENPPFYGVIRWIGOP 487
Db 245 LQELP-----NDPSLGVGSMVEVTLDDGAFAPGSATAPPATMTDPLLYGVIRWIGPL 297
QY 488 P--GLN--EVLAGLELEDE-----CAGCTDGTGTRGYFTFCALKKALFVKLKSRCPSRFA 539
Db 298 PTGGGNHRKVMGVVELEDEPIDFTLETTNGTHNGVLFRCPANRALFVHTSQSRDRRFQ 357
QY 540 SIQPVSNQIERCNSLA-----FGGYLSEVVEENTPPK--MEKEGLEIMIGKKGI 587
Db 358 DIPPMSPCSRTTPAAGSGSKTDNNMFGKVDPCVVRGVRPPLKILKLELEETICGKFKGI 417
QY 588 QGHYNSCYLDSTLFCIFAPSSVLDTVLLRPKEKNDVEYYSETOELLRTIYNPLRIYGV 647
Db 418 QGHNSCYLDATLFAFTFTSVFDSLLFRKPEPNQYEEVQVRLLEIYNPLRNHFV 477
QY 648 CATKMKILKLEKVAASGFTSEEDPREFLNLFPHILVRPLLKIRSAQGVQVDCYF 707
Db 478 RADRVVLKQLDLRLSVTGLMSEKDPFELNLLAQILRADPFLKLSG-----LDTFY 533
QY 708 YQIFMEKNEKVGVTPTQQLEWFSFINSNLKFAEAPSLTIQMPRFGDKFLFKKIPPSLE 767
Db 534 YQLFVEKDERNLPSVQQLFEQSLASNLKLEVPSCLLIQMPRFGKFNKFWYRILPSQV 593
QY 768 LNTDLELTPRCRIQGLAMYTECRECYDDPDISAGTKIQFCKTQNTQVHLHPKRLNHK 827
Db 594 LDVTDIIESPRQVWCWCGKLAIEYECRECFQKMQCEGLEGTAKCSCTDSVHHHSKRLNHK 653
QY 828 YNPVSLPKOL---PDWDWRHGICIPQNMELFAVLCTIETSHYVAFVKYK-SDSAMLFFD 882
Db 654 PVLSPVQDFIPMAP-----HCSVPLRYMELFAVLCIETSHYVAFVKAASGQDAPCFDF 708
QY 883 SMADRDGQGNFPIQVTPCEVGEYL-----KMSLEDLSLDRRIQGCARRLLCDAYMC 938
Db 709 SMADRNGEQNGYNIKPMVPVPDLPRWLTEBGRSALNE-EAVNDKMLPEHAKRLLCDAYMC 767
QY 939 MYQSTPWSLYK 949
Db 768 MYQSTDVMMYR 778
[1]
RESULT 5
Q81PC3 PRELIMINARY; PRT; 551 AA.
ID Q81PC3
AC Q81PC3;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE CG5603-PB.
GN ORFNames=CG5603;
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=107311132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
```

```
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gaber G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Borchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fodor C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush G.F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodgett J., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195 (2000).
RN [2]
SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Tabor P.B., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstein G., Scher S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila
RL melanogaster euchromatic genome sequence.";
RN Genome Biol. 3:RESEARCH0079-RESEARCH0079 (2002).
RN [3]
SEQUENCE FROM N.A.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
RL a genomics perspective.";
RN Genome Biol. 3:RESEARCH0084-RESEARCH0084 (2002).
RN [4]
SEQUENCE FROM N.A.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradscky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RL systematic review.";
RN Genome Biol. 3:RESEARCH0083-RESEARCH0083 (2002).
RN [5]
SEQUENCE FROM N.A.
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
```


RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review";
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN [5]

RG SEQUENCE FROM N.A.
RP FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [6]

RA SEQUENCE FROM N.A.

RG FlyBase;

RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.

DR EMBL; AE003628; AAN10740.1; -

DR FlyBase; FBgn0032210; CG5603.

DR InterPro; IPR000938; CAP-Gly.

DR Pfam; PF01302; CAP_GLY; 1.

SQ SEQUENCE 550 AA; 62083 MW; C87C73651D88FD59 CRC64;

Query Match 22.5%; Score 1134; DB 2; Length 550;
Best Local Similarity 45.7%; Pred. No. 1.4e-65;
Matches 238; Conservative 87; Mismatches 168; Indels 28; Gaps 13;

QY 447 VOESPLAPPNGSHGLEVSLAEVKNP-----PFYGVIRWIGPPG-LNEVLGLGLELED 501
DB 38 IADYQLEIHPGTE--LAIGSLVEV-SNFGVCEDLGVVRWIGIPGPGQKVLVGLIEVED 94
QY 502 EC-----AGCTDGTFRGTRVFTCALKALFKVLKSCRPSRFPASLQP--VSNQIERCNLSA 555
DB 95 ESNLKNVWASDRHNGVRLFTCHDGRAIFVPANRCTADRRFADVDNSISANRVSNNHAK 154
QY 556 FGGYLSSEVVEENTPPKM--EKGLEIMICKKGIQGHNSCYLDSTLFCLPAPSSVLDTV 613
DB 155 FGADCPAIIYSGIPPLQIHNSDELASICGKFGIQQHNSCYLDATLFSMTFTSVFDSI 214
QY 614 LRPKEKNDVEYVSETOELLRTIEIYNPLRIYGVVCAIKMKILKLEKVEAASGFTSEK 673
DB 215 LYRRPGQDIRYSEVQKVRDEIYNPLRNVPVRSDRVKMLRELLDQSSVGLTCEEK 274
QY 674 DPEEFNLILFHILRVEPLKIRSAQKQVQDCYFYQFMKNEKVGVPITQQLLEWSFIN 733
DB 275 DPEEFNLSLSQIMRVEPEPLKSSG---QDSFYQLFVEKEKLTLPVSQLFEQSFS 330
QY 734 SNLKAFAAPSCLIIQMPRGKDFPKPKTIPPSLELNTLLETPQCRI CGSLAWYECR 793
DB 331 SDIKLKEVPSCFIIQMPREGKGYMYPRILPSQLVDVTIENSPPRQCSLGCGLAEYECR 390
QY 794 EGYDDPDISAG-KIQOFCKTCTQVHLHPKLNHKNYPVSLPKDLPDWDWRHCCIFCQNM 852
DB 391 DCFGSLQAGSGLECTAFCPKCLKTHSHIKRTHNVSKKIYSPKEFKIMA-EHNVVPRLYM 449
QY 853 ELFAVLCTIETSHYVAFVKYK--DDSAWLFDDSMADRDGGONGFNIPOVTPCPVEGYLK- 910
DB 450 ELFAVVCITETSHYVAFVKSGSGDAPWCFDSDMADRGQNGYNIPEITCVPELTQWLSE 509
QY 911 ---MSLEDLHLSDRRIQCARLLCDAYNCMYQSTMSLY 948
DB 510 ECARGINET-STNDKVLPEHAKRIFCDAYMCLYQSTDIMY 549

RESULT 8

Q8SYFO PRELIMINARY; PRT; 639 AA.
ID Q8SYFO
AC Q8SYFO
DT 01-JUN-2002 (TremBLrel. 21, Created)
DT 01-JUN-2002 (TremBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TremBLrel. 26, Last annotation update)
DE RE64280p.
GN ORFName=CG5603;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;

RA SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarini H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Ceiniker S.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY071592; AAL49214.1; -

DR FlyBase; FBgn0032210; CG5603.

DR InterPro; IPR000938; CAP-Gly.

DR Pfam; PF01302; CAP_GLY; 1.

SQ SEQUENCE 639 AA; 72162 MW; 6E29D3B8B757675B CRC64;

Query Match 22.5%; Score 1132.5; DB 2; Length 639;
Best Local Similarity 41.2%; Pred. No. 2.2e-65;
Matches 254; Conservative 103; Mismatches 203; Indels 57; Gaps 17;

QY 352 NGSSVDSQPSKSNWYIDVADPAKSLTEI-STDFRSPPLQPPVNSLTENRPH 410
DB 59 NHSADVNDQ-----HLEDV-----DLADILGTNWKRAAGPAAAMLNKSKTD---- 99
QY 411 SLPESLTKMPTNGSICHSPSLSAQSVMEELNTPVQESPPPLAMPNGSHGLEVGLAE 470
DB 100 -----PSNSVDLILKPASPILKTEPEPELRFTIADYQPLEIHPGTE--LAIGSLVE 148
QY 471 VKENP---PFYGVIRWIGPPG-LNEVLGLGLEDEC---AGCTDGTFRGTRVFTCAL 521
DB 149 V-SNPGVCEDLGVVRWIGIPGPGQKVLVGLIEVEDSNLKNVWASDRHNGVRLFTCHD 207
QY 522 KKALFKVLKSCRPSRFPASLQP--VSNQIERCNLSAGFGYLSSEVVEENTPPKM--EKEL 577
DB 208 GRAIFVPANRCTADRRFADVDNSISANRVSNNHAKFGVADCPAIIYSGIPPLQIHNSDEL 267
QY 578 EIMIGKKKGIQGHNSCYLDSTLFCLPAPSSVLDTVLLRPKEKNDVEYVSETOELLRTIEI 637
DB 268 ASICGKFGIQQHNSCYLDATLFSMTFTSVFDSILYRRPGQDIRYSEVQKVRDEI 327
QY 638 VNPRLIYGVVCAIKMKILKLEKVEAASGFTSEKDPPEFLNLFPHILRVEPLKIRS 697
DB 328 VNPRLKNVPVRSDRVKMLRELLDQSSVGLTCEEKDPPEFLNLSQIMRVEPEPLKUSS 387
QY 698 AGOKVQDCYFYQFMKNEKVGVPITQQLLEWSFINSLKFAEAPSCIIQMPRGKDFK 757
DB 388 G---QDSFYQLFVEKEKLTLPVSQLFEQSHSSDIKLEVPSCFIIQMPREGKGYK 443
QY 758 LFKKIPPSLELNTLLETPQCRI CGSLAWYECRECYDDPDISAG-KIQOFCKTCTQV 816
DB 444 MYPRILPSQLVDVTIENSPPRQCSLGCGLAEYECRCFGSLQAGSGLECTAFCPKCLKT 503
QY 817 VHLHPKLNHKNYPVSLPKDLPDWDWRHCCIFCQNMELFAVLCTIETSHYVAFVKYK--DD 875
DB 504 FHSIKRTHNVSKKIYSPKEFKIMA-EHNVVPRLYMELFAVVCITETSHYVAFVKSGSGPD 562
QY 876 SAWLFFDSDMADRDGGONGFNIPOVTPCPVEGYLK---MSLEDLHLSDRRIQCARLL 931
DB 563 APWCFDSDMADRGQNGYNIPEITCVPELTQWFSSEGARSINET-STNDKVLPEHAKRI 621
QY 932 LCDAYNCMYQSTMSLY 948
DB 622 FCDAYMCLYQSTDIMY 638

RESULT 9

Q6TXJ6 PRELIMINARY; PRT; 454 AA.
ID Q6TXJ6
AC Q6TXJ6
DT 05-JUL-2004 (TremBLrel. 27, Created)
DT 05-JUL-2004 (TremBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TremBLrel. 27, Last annotation update)
DE LRRGT00003.

OS	Rattus norvegicus (Rat).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX	NCBI_TaxID=10116;
[1]	
RP	SEQUENCE FROM N.A.
RC	STRAIN=Sprague-Dawley;
RA	Xu C.S., Chang C.F., Han H.P., Wang G.P., Chai L.Q., Yuan J.Y.,
RA	Yang K.J., Zhao L.F., Ma H., Wang L., Wang S.F., Xing X.K., Shen G.M.,
RA	Shi J.B., Rahman S., Wang Q.N., Zhang J.B.;
RL	Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC	-!- SIMILARITY: Belongs to the SIAE family of ribosomal proteins.
DR	EMBL; AY383638; AAQ96216.1; -
DR	GO; GO:0005840; C:ribosome; IEA.
DR	GO; GO:0003735; F:structural constituent of ribosome; IEA.
DR	GO; GO:0006412; P:protein biosynthesis; IEA.
DR	InterPro; IPR001593; Ribosomal_SIAE.
DR	Pfam; PF01015; Ribosomal_SIAE; 1.
DR	ProDom; PD003035; Ribosomal_SIAE; 1.
KW	Ribonucleoprotein; Ribosomal protein.
SQ	SEQUENCE 454 AA; 52009 MW; 8828E384B9F8F69F CRC64;
Query Match 21.7%; Score 1092.5; DB 2; Length 454;	
Best Local Similarity 53.9%; Pred. No. 5.5e-63;	
Matches 213; Conservative 3; Mismatches 2; Indels 177; Gaps 2;	
Qy	555 AFGGYLSVWEZENTPPKMEKEGLEIMICKKGIOGHYNSCYLDSTLFCLFAFSVLDTVL 614
Db	237 AFGGYLSVWEZENTPPKMEKEGLEIMICKKGIOGHYNSCYLDSTLF----- 283
Qy	615 LRPKEKDVEYYSETQELLRTIEVNPLRIYGVCATKMILKILEKVEAASGFTSEKD 674 284 -----
Db	675 PEEFNILPHILRVPELLKTSACQKVDCFYQIFMEKNEKVGVPITIQOLLEWSFINS 734 284 -----W-----
Qy	735 NLKFAEAPSLIIQMFRFGDKFLPKFTPSLELNITDLLEDTPQCRI CGGLAMYECE 794 285 -----QCRICGGLAMYECE 299
Qy	795 CYDDPDISAGIKI KQCKTCNTQVHLHPKRLNHKNYPVSLPKDLPDWWRHGCI PQNMEL 854 ProDom; PD003035; Ribosomal_SIAE; 1.
Db	300 CYDDPDISAGIKI KQCKTKTQTQVHLHPRNLNTHYPVSLPKDLPDWWRHGCI PQNMEL 359
Qy	855 FAVLCIETSHVAFVKY GKDSAMLFFDSMADRDGGQGNGFINIPQVTPCFGEVYLKMSLE 914
Db	360 FAVLCIETSHVAFVKY GKDSAMLFFDSMADRDGGQGNGFINIPQVTPCFGEVYLKMSLE 419
Qy	915 DLHSLSRRIOGCARRLLCDAYMCYQSPTMSLYK 949
Db	420 DLHSLSRRIOGCARRLLCDAYMCYQSPTMSLYK 454
RESULT 10	
Q9VL04	
ID	Q9VL04 PRELIMINARY; PRT; 517 AA.
AC	Q9VL04;
DT	01-MAY-2000 (TREMBLrel. 13, Created)
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT	01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE	CGS603-PA
GN	ORFNames=CGS603;
OS	Drosophila melanogaster (Fruit fly).
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC	Ephydroidea; Drosophilidae; Drosophila.
OX	NCBI_TaxID=7227;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA	Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
 RA April J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballieu R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Rurtis K.C., Buesam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., de
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fouts C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
 RA Jallil M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.C., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Nelson S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nussent D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinartz K., Remington K., Saunders R.D., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodagef, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong M., Zhou X., Zhu S., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RT Science 287:2185-2195(2000).
 RP [2]
 RP SEQUENCE FROM N.A.
 RP MEDLINE=22426065; PubMed=12537568;
 RX Celniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
 RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,
 RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
 RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
 RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
 RT "Finishing a whole-genome shotgun: Release 3 of the *Drosophila*
 RT melanogaster euchromatic genome sequence.";
 RT Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
 RP [3]
 RP SEQUENCE FROM N.A.
 RP MEDLINE=22426070; PubMed=12537573;
 RX Kaminker J.S., Bergman C.M., Krommiller B., Carlson J., Svirskas R.,
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
 RA Ashburner M., Celniker S.E.;
 RT "The transposable elements of the *Drosophila melanogaster* euchromatin:
 RT a genomics perspective";
 RT Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
 RP [4]
 RP SEQUENCE FROM N.A.
 RP MEDLINE=22426069; PubMed=12537572;
 RX Mixra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a
 RT systematic review";
 RT Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
 RP [5]
 RP SEQUENCE FROM N.A.


```
Db 371 KRSDDL-----LNLVCLDLFTKEKLD---VWYKQCHDNRSRAKSLRIKIPP 418
QY 801 ISAGTKIQKCTCNTQVHLHPKRLNHNKYNPVSUPKDLPL-----DWDWR 843
Db 419 VLSIQKRF-----KHNHL-----NDTFSKIEIPIDIPFLNITKYTSASMSPELDCD-- 466
QY 844 HGCIPQNWELFAVL-----IETSHVAVFVKYKDDSAWLPFD 882
Db 467 -----KIFELFALVCHIGSVNTGHVIVIKNG--NQWMLKFD 501

RESULT 14
Q6P5Z9 PRELIMINARY; PRT; 653 AA.
AC Q6P5Z9
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE RSN protein (Fragment).
GN Name=RSN;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.243603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner K.H., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T.J., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ussin T.B., Tohyuki S., Carninci P., Prange C.,
RA Raha S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph;
RA Strausberg R.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC062543; AAH62543.1; -.
DR InterPro; IPR000938; CAP-Gly.
DR Pfam; PF01302; CAP_GLY_2.
DR PROSITE; PS00845; CAP_GLY_1; 2.
DR PROSITE; PS0245; CAP_GLY_2; 2.
FT NON_TER 653
SQ SEQUENCE 653 AA; 71733 MW; 4074A22377D4BB17 CRC64;

Query Match 3.0%; Score 151; DB 2; Length 653;
Best Local Similarity 20.8%; Pred. No. 0.39;
Matches 99; Conservative 55; Mismatches 171; Indels 157; Gaps 20;

QY 260 GYFVGVDMPNIGWDRFGVLCSFACVBSSTILLHNDIIPESVTQERRPPKLAFWSRG 319
Db 85 GQWAGIVLDEPIGKNDGVAGVR-YQCEP-----LKGF-----TRPSKL---TRX 127
QY 320 VGDGKSSHNKPKATGSTDPGNRRSELFYTLNGSSVDSPQSKSNWTYIDEVAEDPAK 379
Db 128 V-----QAEDEANGLQTTASRATSPLCSTASWSSSPSPSN---IPQKPSQPA- 175
```

```
QY 380 SLTEISTDFDRSSPPLOPPPVNSLTNTENPHSLPSPSILTKMNTNGSIHSPLSLSAQSV 439
Db 176 -----AKEPSATPISNLT-----KTASESIS 197
QY 440 EELNTAPVQSPPLAMPNGSHGLEVSLAEVKENPPFYGVIRWIGOPPLNEVLAGLEL 499
Db 198 NLSEAGSIK-----CERELKIGDRVLVGTKA--GVVRLGETDFPAKGEWCGBEL 246
QY 500 EDECAGCTDGTFRGTFTYFCALKKALFKVLKSCRPDSRFASLQPVSNQIERNSLAFGGY 559
Db 247 -DEPLGKNDGAVAGTRYFQCPKYGLFA-----PVHKVTIKGF-- 283
QY 560 LSEVVENTPPKMEKGL-PIMIGKKGIQHNVCYLDSTLFLCFAPSSVLDTVLLRPK 618
Db 284 -----PSTTPAKAKANAVRVMATTSASLKSPSSASSLS--MSSVASSVSSRP- 330
QY 619 EKNDVEYSETQELLRTIENVLRIVGVCAIKMK-----LRKILEKVEAASGF 668
Db 331 -----SRTGLTETSSRYARKISGTTALQELAKKQKHIEQLLAERDLERAFAKA- 381
QY 669 TSEKQPEEFNLIL-----FHHIL-----RVEPLLKIRSAGQKQVDC 705
Db 382 TSHVGETIEQELALARDGHDQVLELEAKMDQLRTMVEADREKVELLNQLSEKRVDEL 441
QY 706 YF 707
Db 442 QF 443

RESULT 15
REST HUMAN
ID REST HUMAN STANDARD; PRT; 1427 AA.
AC P30622;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 03-JUL-2004 (Rel. 44, Last annotation update)
DE Restin (Cytoplasmic linker protein-170 alpha-2) (CLIP-170) (Reed-
DE Sternberg intermediate filament associated protein).
GN Name=RSN;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Peripheral blood monocytes;
RX MEDLINE=92289675; PubMed=1600942;
RA Bilbe G., Delabie J., Brueggen J., Richener H., Asselbergs F.A.M.,
RA Cerletti N., Sorg C., Odink K., Tarcsay L., Wiesendanger W.,
RA de Wolf-Peeters C., Shipman R.;
RT "Restin: a novel intermediate filament-associated protein highly
RT expressed in the Reed-Sternberg cells of Hodgkin's disease."
RL EMBO J. 11:2103-2113(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92405160; PubMed=1356075; DOI=10.1016/0092-8674(92)90240-D;
RA Pierre P., Scheel J., Rickard J.E., Kreis T.E.;
RT "CLIP-170 links endocytic vesicles to microtubules."
RL Cell 70:887-900(1992).
CC -I- FUNCTION: Seems to be a intermediate filament associated protein
CC that links endocytic vesicles to microtubules.
CC -I- SUBCELLULAR LOCATION: Cytoplasmic; associated with the
CC cytoskeleton.
CC -I- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Name=Long;
CC IsoId=P30622-1; Sequence=Displayed;
CC Name=Short;
CC IsoId=P30622-2; Sequence=VSP_000765;
CC -I- TISSUE SPECIFICITY: Highly expressed in the Reed-Sternberg cells
CC of Hodgkin's disease.
CC -I- SIMILARITY: Contains 2 CAP-Gly domains.
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 18, 2005, 15:25:35 ; Search time 1059 Seconds
(without alignments)
5304.848 Million cell updates/sec

Title: US-09-671-687A-3

Perfect score: 5034

Sequence: 1 MSGGLWSQKVTSPWEERI.....RLLCDAYMCVQSPTMSLYK 949

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DBV=xlp
-Q/cgn2_1/USPTO.spool_p/US09671687/runat.18042005.115017.17115/app_query.fasta_1.1095
-DB=N_Geneseq_16Dec04 -QFW=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -ENDS=1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09671687 @CGN 1.1.708 @runat.18042005.115017.17115 -NCPU=6 -ICPU=3
-NO_WRAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq_16Dec04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4983	99.0	3302	12	Adq95917 T cell ac
2	4983	99.0	6831	10	Adc24898 Human bre
3	4980	99.9	3302	12	Adq95915 T cell ac
4	4971.5	98.8	5371	10	Abv75394 Human CYL
5	4971.5	98.8	5371	13	Adrl4488 Human NF-

6	4968.5	98.7	3311	12	Adq95919	Adq95919 T cell ac
7	4942	98.2	4527	3	AAa08589	AAa08589 Human cyt
8	4601	91.4	2845	4	AAh18625	AAh18625 Human cdn
9	3876.5	77.0	4716	4	AAa41617	AAa41617 CDNA enco
10	3676	73.0	2523	4	AAi58410	AAi58410 Human pol
11	3676	73.0	2523	5	Adq98620	Adq98620 DNA enco
12	3676	73.0	2523	9	AdA48380	AdA48380 Novel hum
13	3523	70.0	4286	4	AAi160196	AAi160196 Human pol
14	3329	66.1	2488	4	AAa41081	AAa41081 CDNA enco
15	3022.5	60.0	2569	4	AAH18478	AAH18478 Human cdn
16	2468	49.0	2116	2	AAZ07515	AAZ07515 Human RAP
17	1778	35.3	1151	6	ABL89642	ABL89642 Human pol
18	1494	29.7	1013	6	ABL89643	ABL89643 Human pol
c	1490.5	29.6	1160	5	AAa90524	AAa90524 DNA enco
c	1262	25.1	837	4	AAH07749	AAH07749 Human cdn
20	1215	24.1	758	4	AAH05084	AAH05084 Human cdn
c	1102	21.9	617	6	ABQ60558	ABQ60558 Human col
c	936	18.6	5433	4	ABL09704	ABL09704 Drosophil
c	879	17.5	557	6	ABQ60559	ABQ60559 Human col
25	867.5	17.2	8990	4	AAK77783	AAK77783 Human imm
26	856	17.0	1655	4	ABL09705	ABL09705 Drosophil
27	626	12.4	376	8	ABX51704	ABX51704 Bovine ES
28	568	11.3	318	12	ADK71948	Adk71948 Human ori
29	545	10.8	303	12	ADK71946	Adk71946 Human ori
c	545	10.8	309	4	AAI20692	AAI20692 Probe #10
c	545	10.8	309	4	ABA65741	ABA65741 Human foe
c	545	10.8	309	4	AAI45907	AAI45907 Probe #14
c	545	10.8	309	4	ABA47848	ABA47848 Human bre
c	545	10.8	309	4	ABA32825	ABA32825 Probe #11
c	545	10.8	309	4	AAK39883	AAK39883 Human bon
c	545	10.8	309	4	AAK14143	AAK14143 Human bra
c	545	10.8	309	5	AAI06390	AAI06390 Probe #63
c	545	10.8	309	6	ABS13981	ABS13981 Human gen
c	514	10.2	288	12	ADK71950	Adk71950 Human ori
c	512	10.2	483	4	AAI11479	AAI11479 Probe #14
c	512	10.2	483	4	ABA53149	ABA53149 Human foe
c	512	10.2	483	4	AAI32754	AAI32754 Probe #14
c	512	10.2	483	4	ABA42721	ABA42721 Human bre
c	512	10.2	483	4	ABA22920	ABA22920 Probe #13
c	512	10.2	483	4	AAK26850	AAK26850 Human bon

ALIGNMENTS

RESULT 1
ADQ95917
ID ADQ95917 standard; cdna; 3302 BP.
XX
AC ADQ95917;
XX
DT 07-OCT-2004 (first entry)
XX
DE T cell activation associated cdna #48.

ss; gene; antiallergic; antiarthritic; antidiabetic;
anti-HIV; antimicrobial; antiarthritic; immunosuppressive;
neuroprotective; gene therapy; T cell activation; diagnosis;
autoimmune disease; rheumatoid arthritis; asthma; multiple sclerosis;
diabetes; allergic disease; infectious disease; AIDS; chronic rejection;
organ; bone-marrow transplant.
Homo sapiens.

Key Location/Qualifiers
CDS 243..3104
/*tag= a

WO2004058805-A2.
15-JUL-2004. - *late*
25-DEC-2003; 2003WO-JP016715.

QY 517 PheThrCysAlaLeuLysAlaLeuPheValLysLeuLysSerCysArgProAspSer 536
Db 1803 TTCACCTGTCCTGAAGAAGCGCTGTTGTGAAACTGAAGAGCTGCAGCGCTGACTCT 1862
QY 537 ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe 556
Db 1863 AGTTTGCACTCATGAGCGGTTTCATCAGATTGAGCGCTGTAACTCTTTAGCATTT 1922
QY 557 GlyGlyTyrLeuSerGluValValGluAsnThrProProLysMetGluLysGluGly 576
Db 1923 CGAGGCTACTTAAGTCAAGTAGTAGAAGAAATFACCCACAAAATGGAAAAAAGAGGC 1982
QY 577 LeuGluIleMetIleGlyLysLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu 596
Db 1983 TTGGAGATAATGATTGGGAAGAGAAAGGCATCCAGGGTCATTACAAATCTCTGTACTTA 2042
QY 597 AspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArg 616
Db 2043 GACTCAACCTTATCTGCTTATTTGCTTTAGTTCTGTCTGGACACTGTGTACTTAGA 2102
QY 617 ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu 636
Db 2103 CCCAAAGAAAGAACGATGAGATATATAGTGAACCCAGAGCTACTGAGACAGAA 2162
QY 637 IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg 656
Db 2163 ATTGTTAATCCTCTGAGAATATATGATATGTGTGCCACAAAAATTTATGAACCTGAGG 2222
QY 657 LysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGlu 676
Db 2223 AAAATACTTTGAAAAGGTGGAGGCTGATCAGGATTTTACCTCTCAAGAAAAGATCTCGAG 2282
QY 677 GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLysIleArg 696
Db 2283 GAATCTTGAATATCTGTTTCATCATATTTTAAAGGGTAGAACCCTTTGCTAAAAATAGA 2342
QY 697 SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu 716
Db 2343 TCAGCAGGTCAAAAGGTACAAGATTGTACTTCTATCAAAATTTTATGGAATAAATGAG 2402
QY 717 LysValGlyValProThrIleGlnGlnLeuLeuGluTyrSerPheIleAsnSerAsnLeu 736
Db 2403 AAAGTTGGCGTTCCCAATTCACAGATTGTTTGAATGGTCTTTTATCAACAGTAACCTG 2462
QY 737 LysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPhe 756
Db 2463 AAATTTGCAGAGCCACCATCATGCTGATTATTCAGATGCTCGATTGGAAAAGACTTT 2522
QY 757 LysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAsp 776
Db 2523 AAATATTAAAAAATTTTCTCTCTCGAATTAATAATATACAGATTACTTGAAGAC 2582
QY 777 ThrProArgGlnCysArgIleCysGlyLeuAlaMetTyrGluCysArgGluCysTyr 796
Db 2583 ACTCCAGACAGTCCGCGATATGCGAGGCTTGCAATGTATGAGTGTAGAGAACTGAC 2642
QY 797 AspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGln 816
Db 2643 GACGATCCGGACATCTCAGCTGGAAAATCAACAGTTTTTGTAAAAACCTGCAACACTCAA 2702
QY 817 ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAsp 836
Db 2703 GTCCACTTCATCCGAAGAGGCTGATCATATAATATAACCATGTGCTACTTCCCAAGAC 2762
QY 837 LeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAla 856
Db 2763 TTACCCGACTGGAGTGGACACGCGTGCATCCCTTGCAGAAATATGAGATTATTGCT 2822
QY 857 ValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSer 876
Db 2823 GTTCTCTGATAGAAAACGCACTATGTTGCTTTTGTGAAGTATGGGAAGGACGATTCT 2882

QY 877 AlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIle 896
Db 2883 GCTGCTCTCTTTTACAGCATGGCGATCGGATGGTGGTCAATATGGCTTCAACAT 2942
QY 897 ProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLeu 916
Db 2943 CCTCAAGTCAACCCCATGCCAGAAAGTAGAGAGTACTTTGAAGATGTCTCTGGAAGACCTG 3002
QY 917 HisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyr 936
Db 3003 CATTCCTTGGACTCCAGGAATCCAAAGGCTGTGCACGAAGACTGCTTTGTGATGCATAT 3062
QY 937 MetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
Db 3063 ATGTGCATGTACCAGAGTCCAAATGAGTTTGTACAAA 3101
RESULT 2
ADC24898
ID ADC24898 standard; cDNA; 6831 BP.
XX
AC ADC24898;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human breast specific nucleic acid (BSNA) DEX0238_49, SEQ ID NO:49.
XX
KW Human; breast specific polypeptide; BSP; breast specific nucleic acid;
KW BSNA; breast cancer; metastasis; non-cancerous disease; breast tissue;
KW identification; monitoring; diagnosis;
KW engineered breast tissue production; transgenic animal; drug screening;
KW cytostatic; gene therapy; vaccine; chromosome 16p13.3; gene; ss.
XX
OS Homo sapiens.
XX
PN WC02003020900-A2.
XX
PD 13-MAR-2003.
XX
PF 29-AUG-2002; 2002WO-US027777.
XX
PR 31-AUG-2001; 2001US-0316306P.
XX
PA (DIAD-) DIADEXUS INC.
XX
PI Sun Y, Liu C, Salceda S;
XX
DR WPI; 2003-290182/28.
DR P-PSDB; ADC24816.
XX
PT New breast specific polypeptide useful for identifying, diagnosing,
PT monitoring, staging, imaging and treating breast cancer and non-cancerous
PT disease states in breast.
XX
PS Example 1; SEQ ID NO 49; 264pp; English.
XX
CC The invention relates to breast specific polypeptides (BSNs) and nucleic
CC acids (breast specific nucleic acids; BSNAs) encoding them. The invention
CC also relates to vectors and host cells comprising a BSNA sequence;
CC antibodies against BSNs; the recombinant production of BSNs; methods of
CC detection of BSNAs or BSNs in a sample; kits for detecting a risk of
CC cancer or presence of cancer in a patient; and vaccines comprising a BSNA
CC or BSP. The invention additionally discloses fragments, mutants, fusion
CC proteins, homologous proteins and allelic variants of BSNs; methods for
CC identifying and designing agonists and antagonists of BSNs; methods for
CC identifying and monitoring breast tissue; producing engineered breast
CC tissue for treatment and research; producing transgenic animals and cells
CC comprising BSNA sequences; aptamers evolved to bind specifically to BSNs;
CC and single exon probes based on BSNA sequences. BSNs, BSNAs and
CC antibodies against BSNs are useful for identifying, diagnosing,
CC monitoring, staging, imaging and treating breast cancer (including breast
CC cancer metastases), and non-cancerous disease states in breast tissue.
CC BSNs and BSNAs may additionally be used to identify and monitor breast
CC tissue, in screening for BSP agonists and antagonists, and in the

CC production of engineered breast tissue for treatment or research. BSNAs
 CC may also be used in gene therapy and in the production of transgenic
 CC animals and cells. The present sequence represents a breast specific
 CC nucleic acid (BSNA) used in an example of the invention.

XX
 SQ Sequence 6831 BP; 1987 A; 1256 C; 1494 G; 2094 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0 Length: 6831
 Score: 4983.00 Matches: 948
 Percent Similarity: 99.48% Conservative: 0
 Best Local Similarity: 99.48% Mismatches: 1
 Query Match: 98.99% Indels: 4
 DB: 10 Gaps: 4

US-09-671-687A-3 (1-949) x ADC24898 (1-6831)

QY 1 MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluGluArgIle 20
 DB 1585 ATGAGTTTCAGGCTTATGGAGCCCAAGAAAAAGTCACCTTCCCTTACTGGGAAGAGCGGATT 1644
 QY 21 PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLeuLys 40
 DB 1645 TTTTACTTCTCTTCAAGAATCCAGCGTTACAGACAAACAAACAAAGCTCCTTAAA 1704
 QY 41 ValProLysGlySerIleGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60
 DB 1705 GTACCGAAGGAGTAGTACAGACAGTATATTCAGATCGTTCCTGGGGCATTCAGGATT 1764
 QY 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGlnProHisAla 80
 DB 1765 CCTTCTGCAAAAGGCAAGAAAAATCAGATTGGAATTAATAATTTAGAGCAACCTCATGCA 1824
 QY 81 ValLeuPheValAspGlu---AspValValGluLeuAsnGluLysPheThrGluLeuLeu 99
 DB 1825 GTTCTCTTTGATGAAAGGATGTTGTAGAGATAAATGAAAGTTTCACAGAGTTACTT 1884
 QY 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
 DB 1885 TTGGCAATTACCAGTTGTGAGGAGAGGTTTCAGCTGTTTAAACACAGAACAGACTAAGT 1944
 QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
 DB 1945 AAAGGCTTCCAAATAGACGTGGGCTGTCTGTGAAAGTACAGCTGAGATCTGGGGAGAA 2004
 QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
 DB 2005 AAATTTCTCGAGTTGTACGCTTTCAGAGGACCCCTGTTCAGCAGAGGACAGTCTCCGGA 2064
 QY 160 IlePhePheGlyValGluLeuLeuGluGluGlyArgGlyGlnGlyPheThrAspGlyVal 179
 DB 2065 ATATTTCTTGGAGTTGATTTGCTGGAAAGAGTTCGTTGCTCAAGTTTCACTGACGGGGTG 2124
 QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
 DB 2125 TACCAAGGGAACACAGCTTTTTCAGTGTGATGAAGATTGTGGCTGTTTGTTCATTGGAC 2184
 QY 199 LysLeuGluLeuLeuGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
 DB 2185 AACCTAGAACTCATAAGAGATGATGACACTGCACTTGGAAAGTATTACGCAAGTCTCTGGG 2244
 QY 219 AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly 238
 DB 2245 GACACAAATGACGTCGACATCTCTCTTGGAAATAAATCCAGAGTTTCTTTGAAGTT 2304
 QY 239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
 DB 2305 GGAGAAACAATAGAACTCGAAACAGTTATATCTGTGATGTTTGGCCAGGAAAGAAAGC 2364
 QY 259 LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe 278
 DB 2365 TTAGGATATTTTGTGTGTGGACATGGATAACCCCTATTGGCACTGGGATGGAAGATT 2424

QY 279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
 DB 2425 GATGGAGTCAGCTTTGTAGTTTGGTGTGTGAAAGTACAAATCTATTATTCACATCAAT 2484
 QY 298 AspIleIleProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMetSer 317
 DB 2485 GATATCATCCAGAGAGTGTGACGAGGAAGAGGCGCTCCCAAACTTGGCTTTATGTCA 2544
 QY 318 ArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThrGlySerThr 337
 DB 2545 AGAGGTGTGGGGACAAAGTTTCATCCAGTCATAATAAACCAAGGCTACAGGATCTACC 2604
 QY 338 SerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVal 356
 DB 2605 TCAGACCTTGGAAATAGAAACACAGATCTGAATATTTTATACCTTAAATGGGTCTCTGTT 2664
 QY 357 AspSerGlnProGlnSerLysSerLysAsnThrTyrIleAspGluValAlaGluAsp 376
 DB 2665 GACTCAACACCAATCCAAATCAAAATAATACATGTGTACATTTGATGAAGTTGCAGAAAGC 2724
 QY 377 ProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerProProLeuGln 396
 DB 2725 CTGCAAAATCTTACAGAGATATCTACAGACTTTGACCGTCTTCCACCACCTCCAG 2784
 QY 397 ProProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeu 416
 DB 2785 CCTCTCTCTGTGAACCTCACTGACCCAGAGAACAGATTCACCTCTTTACCATTCAGTCTC 2844
 QY 417 ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGln 436
 DB 2845 ACCAAGATCCCAATACCAATGGAAGTATTGGCCACAGTCCACTTCTCTGTGAGCCAG 2904
 QY 437 SerValMetGluGluLeuLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPro 456
 DB 2905 TCTGTATGGAAGAGCTAAACACTGCACCCGCTCAAGAGAGTCCACCTTGGCCATCGCT 2964
 QY 457 ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro 476
 DB 2965 CCTGGGAACCTCATGCTCTAGAAGTGGGCTCATTTGGCTGAAGTTAAGAGAACCTCTCT 3024
 QY 477 PheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGly 496
 DB 3025 TTCTATGGGGTAATCCGTTGGATCGGTGACCCACAGGACTGAATGAAGTGTCTCGCTGGA 3084
 QY 497 LeuGluLeuLeuAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr 516
 DB 3085 CTGGAACCTGGAAGATGAGTGTGAGCTGTACGATGGAACTTCAGAGGCACTCGGTAT 3144
 QY 517 PheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSer 536
 DB 3145 TTCACCTGTGCTGAAGAAGGCGCTGTTGTGAACTGAAGAGCTGCAGGCTGACTCT 3204
 QY 537 ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe 556
 DB 3205 AGGTTTGCATCATTTGACGCGGTTTCCAATCAGATTGAGCGCTGTAACTCTTTTACGATTT 3264
 QY 557 GlyGlyTyrLeuSerGluValValGluAsnThrProProLysMetGluLysGluGly 576
 DB 3265 GGAGGCTACTTAAGTAGAGTAGAAGAAATACTCCACAAANATGGAAAAAGAGGC 3324
 QY 577 LeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu 596
 DB 3325 TTGGAGATATGATGGGAAGAAAGAGCATCCAGGCTCATTAACAATTTCTTGTACTTA 3384
 QY 597 AspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArg 616
 DB 3385 GACTCAACCTTATTCTGCTTATTGCTTTTGTAGTTCTGTTCTGGACACTGTGTTACTTGA 3444
 QY 617 ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu 636
 DB 3445 CCAGAGAAAGAACGATGATGATATTTATGTGAAACCCAGAGACTACTGAGGACAGAA 3504
 QY 637 IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg 656

Db 3505 ATTGTTAATCTCTGAGATATATGATATGTTGTCACAAAATTTATGAACCTGAGG 3564
 Qy 657 LysIleLeuGluLysValGluAlaSerGlyPheThrSerGluGluLysAspProGlu 676
 Db 3565 AAAATACTTTGAAAGGTGGAGGCTGATCATCAGGATTTACCTCTGAAGAAAAGATCTCTGAG 3624
 Qy 677 GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArg 696
 Db 3625 GAATCTTGAAATATCTGTTTCATCATATATTTAAGGGTAGAACCTTTGCTAAAATAAAGA 3684
 Qy 697 SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu 716
 Db 3685 TCAGCAGGTCAAAAGGTACAGATTTGTTACTTCTATCAAAATTTTATGGAAAAATGAG 3744
 Qy 717 LysValGlyValProThrIleGlnLysLeuLeuGluTyrPheIleAsnSerAsnLeu 736
 Db 3745 AAAGTTGGGGCTTCCCAATTCAGCAAGTCTGTAGAAATGGTCTTTTATCAACAGTAACCTG 3804
 Qy 737 LysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPhe 756
 Db 3805 AAATTTGCAGAGGCACCATCATCTCGATTTATTCAGATGCTCGATTTGGAAAAGACTTT 3864
 Qy 757 LysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAsp 776
 Db 3865 AAATTAATTTAAATAATTTTCTCTCTGGAATTAATAATAACAGATTTACTTTGAAGAC 3924
 Qy 777 ThrProArgGlnCysArgIleCysGlyLeuAlaMetTyrGluCysArgGluCysTyr 796
 Db 3925 ACTCCAGACAGTGGCGGATATCTGGAGGGCTTGGCAATGTATGAGTGTAGAGAATGCTAC 3984
 Qy 797 AspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGln 816
 Db 3985 GACGATCCGACATCTCAGCTGGAAAATCAACAGATTTTGTAAACCTGCAACACTCAA 4044
 Qy 817 ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAsp 836
 Db 4045 GTCCACCTTATCCGAAGAGGCTGGAATCATAAATAAATCAACAGTGTCCCAAGAC 4104
 Qy 837 LeuProAspTyrAspTyrArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAla 856
 Db 4105 TTACCCGACTGGAGTGGAGACAGGCTGCATCCCTTGCAGAAATATGGATTTATTTGCT 4164
 Qy 857 ValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspSer 876
 Db 4165 GTTCTCTGCATAGAAAACAGCCACTATGTCTTTGTGAAGTATGGGAAGGACGATCT 4224
 Qy 877 AlaTyrLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIle 896
 Db 4225 GCCTGGCTCTCTTTGACAGCATGGCGGATCGGGATGGTGCAGAAATGGCTTCAACATT 4284
 Qy 897 ProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLeu 916
 Db 4285 CCTCAAGTCAACCCATGCCAGAGTAGGAGATCTTTGAAGATGCTCTGGGAAGACCTG 4344
 Qy 917 HisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyr 936
 Db 4345 CATTCCTTGACTCCAGGAGATCCAGAGCTGTGCACGAGAGACTGCTTTGTGTGATAT 4404
 Qy 937 MetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
 Db 4405 ATGTGCATGTACCAAGTCCAAACATCAGTTGTACAAA 4443
 RESULT 3
 ADQ95915
 ID ADQ95915 standard; cDNA; 3302 BP.
 XX
 AC ADQ95915;
 XX
 AC ADQ95915;
 XX
 DT 07-OCT-2004 (first entry)
 XX
 DE T cell activation associated cDNA #47.
 XX

ss; gene; anti-allergic; antiarthritic; antiasthmatic; antidiabetic;
 anti-HIV; antimicrobial; antirheumatic; immunosuppressive;
 neuroprotective; gene therapy; T cell activation; diagnosis;
 autoimmune disease; rheumatoid arthritis; asthma; multiple sclerosis;
 diabetes; allergic disease; infectious disease; AIDS; chronic rejection;
 organ; bone-marrow transplant.

OS Homo sapiens.
 XX
 XX
 PH Key Location/Qualifiers
 CDS 243..3104
 /*tag= a

W02004058805-A2.

15-JUL-2004.

25-DEC-2003; 2003WO-JP016715.

26-DEC-2002; 2002JP-00376365.

27-DEC-2002; 2002US-0436473P.

25-APR-2003; 2003JP-00122113.

28-APR-2003; 2003US-0465792P.

21-OCT-2003; 2003JP-00360559.

22-OCT-2003; 2003US-0512846P.

(ASAH-) ASahi KASEI PHARMA CORP.

Matsuda A, Yoneta S;

WPI; 2004-593134/57.

P-PSDB; ADQ95916.

New purified protein involved in T cell activation, useful for
 diagnosing, preventing and/or treating acquired immunodeficiency
 syndrome, autoimmune (e.g. rheumatoid arthritis, and diabetes), allergic
 and infectious diseases.

Claim 4; SEQ ID NO 93; 2828bp; English.

The invention relates to purified proteins and genes encoding them, that
 are involved in T cell activation (1) and has an amino acid deletion,
 substitution or addition in the amino acid sequences. The methods and
 compositions of the present invention are useful for the diagnosis,
 prevention and/or treatment of autoimmune disease (rheumatoid arthritis,
 asthma, multiple sclerosis and diabetes), allergic disease, infectious
 disease, AIDS, and acute or chronic rejection at organ transplant or bone
 marrow transplant. This sequence corresponds to a cDNA for a protein
 involved in T cell activation.

Sequence 3302 BP; 981 A; 648 C; 772 G; 901 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0 Length: 3302
 Score: 4980.00 Matches: 947
 Percent Similarity: 99.48% Conservative: 1
 Best Local Similarity: 99.37% Mismatches: 1
 Query Match: 98.93% Indels: 4
 DB: 12 Gaps: 4

US-09-671-687A-3 (1-949) x ADQ95915 (1-3302)

Qy 1 MetSerSerGlyLeuTyrSerGlnLysValThrSerProTyrTyrGluArgIle 20
 Db 243 ATGAGTTTCAGCTTATGGAGCCAGAAAAGTCACTTCACTTCTGGGAAGACGGATT 302
 Qy 21 PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLys 40
 Db 303 TTTTACTTGCTTCTTCAAGATGCGAGTTACACACAAACACAAACAGCTCTTAAA 362
 Qy 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60
 Db 363 GTACCCGAAGGGAAGTATAGGACAGTATATTCAAGATCGTTCTGTGGGCGCATTCAGGATT 422

QY	61	ProSerAlaIysGlyLysLysGlnIleGlyLeuLysIleLeuGluGlnProHisAla	80	QY	417	ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGln	436
Db	423	CCTTCTGCAAGGCAAGAAATCAGATTGGATTAAAAATTCAGACCAACCTCATGCA	482	Db	1503	ACCAAGATGCCAATACCAATGGAGTATTGGCCACAGTCCACTTCTCTGTGAGCCAG	1562
QY	81	ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu	99	QY	437	SerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPro	456
Db	483	GTCTCTTTTGTGATGAAAGGATGTTGAGAGATAAATGAAAGTTTCACAGAGTTACTT	542	Db	1563	TTCTGTAATGGAGAGCTAAACACTGCCCGTCCAAGAGAGTCAACCTTGGCCATGCCT	1622
QY	100	LeuAlaIleThrAsnGlyGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer	119	QY	457	ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro	476
Db	543	TTGGCAATTTACCAATTTGTGAGGAGGTTTCAGGCTGTTTAAAAACAGAACAGACTAAGT	602	Db	1623	CCTGGGAACCTCACATGCTAGAGTGGGCTCATTTGGCTGAAGTTAAGAGAACCTCTCT	1682
QY	120	LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu	139	QY	477	PheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGly	496
Db	603	AAAGGCTCCAAATAGACGTGGGCTCTCTGTGAAAGTACAGCTGAGATCTGGGGAAGAA	662	Db	1683	TTCTATGGGGTAAATCCGTTGGATCGGTCAGCCACAGGACTGAATGAAGTCTCGCTGGA	1742
QY	140	LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly	159	QY	497	LeuGluLeuLeuAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr	516
Db	663	AAATTTCTGGAGTTGTACGCTTTCAGAGACCCCTGTTAGCAGAGGACAGTCTCCGGA	722	Db	1743	CTGGAACTGGAAGATGAGTGTGAGGCTGTACGATGGAAACCTTCAGAGGCACTCGGTAT	1802
QY	160	IlePhePheGlyValGluLeuLeuGluGluArgGlyGlnGlyPheThrAspGlyVal	179	QY	517	PheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSer	536
Db	723	ATATTTCTTGGAGTTGAAATGCTGGAGAGAGTCTGGTCAAGGTTTCACTGACGGGGT	782	Db	1803	TTACCTGTGCTGAGAGGCGCTGTTGTGAACTGAAGAGCTGCAGGCTGACTCT	1862
QY	180	TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp	198	QY	537	ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe	556
Db	783	TACCAAGGAAACAGCTTTTTCAGTGTGATGAGATTGTGGCGTGTGTTGTCATTGGAC	842	Db	1863	AGGTTTGCATCATTTGACGCGGTTTCCAATCAGATTGAGCGCTGTAATCTCTTACGATTT	1922
QY	199	LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly	218	QY	557	GlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluGly	576
Db	843	NAAGTAGAACTCATAGAGATGATGACACTGCATTCGAAAGTGAATACGCGGCTCTGGG	902	Db	1923	GGAGGCTACTTAAAGTAGTAGAAGAAATACTCCACCAAAATGGAAAAAGAGGC	1982
QY	219	AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly	238	QY	577	LeuGluIleMetIleGlyLysLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu	596
Db	903	GACACAATGCAGTCAAGCTTCTCTCTTGGAAATAAACTCCAGAGTTTCTTTGAAAGTT	962	Db	1983	TTGCAATATATGATTGGGAAGAAAGAGGATCCAGGCTCATTACAAATCTTGTACTTA	2042
QY	239	GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer	258	QY	597	AspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArg	616
Db	963	GGGAAACAATAGAACTGGAACAGATTATATCTGTGATGTTTCCAGAGAAAGAAAGC	1022	Db	2043	GACTCAACCTTATTTCTGCTTATTTGCTTTTAGTTCTGTTCTGGACACTGTGTTACTAGA	2102
QY	259	LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe	278	QY	617	ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu	636
Db	1023	TTAGGATATTTTGTGTGTGGACATGATAACCTTATTGGCAACTGGGATGGAAGATT	1082	Db	2103	CCCAAGAAAGAACCGATAGTAATATATAGTGAAACCCCAAGAGCTACTGAGACAGAA	2162
QY	279	AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn	297	QY	637	IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg	656
Db	1083	GATGGAGTGCAGCTTTGTAGTTTGGCTGTGTTGAAAGTACAAATTTCTATTGCAATCAAT	1142	Db	2163	ATTGTTAATCCTCTGAGATATATGGATATGTGTGCCACAAAATATGAACTGAG	2222
QY	298	AspIleIleProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMetSer	317	QY	657	LysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGlu	676
Db	1143	GATATCATCCAGAGAGTGTGACGAGAAAGAGGCTCCCAAACTTGGCTTTATGTCA	1202	Db	2223	AAAATACTTCAAAAGGTGGAGGCTGCATCAGGATTTTACCTCTGAAGAAAAAGATCCTGAG	2282
QY	318	ArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThrGlySerThr	337	QY	677	GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArg	696
Db	1203	AGAGGTGTGGGACAAAGGTTTCATCCAGTCATAATAAACCAAGGCTACAGGATCTACC	1262	Db	2283	GAATTTCTGAAATATCTGTTTTCATCATATTTTAAAGGGTAGAACCTTTGCTAAAAATAAGA	2342
QY	338	SerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerVal	356	QY	697	SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu	716
Db	1263	TCAGACCTCGAAATAGAAACAGATCTGAATATATTTATACCTTAAATGGGCTCTCTGTT	1322	Db	2343	TCAGCAGGTCAAAGGTACAGGATTTGTTACTTCTATCAAAATTTTATGGAAAAAATGAG	2402
QY	357	AspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGluValAlaGluAsp	376	QY	717	LysValGlyValProThrIleGlnGlnLeuLeuGluTrpSerPheIleAsnSerAsnLeu	736
Db	1323	GACTCAACCACAATCCAAATCAAAAAATACATGGTACATTGATGAAGTTGCAAGAGAC	1382	Db	2403	AAAGTTGGCGTTCACCAATTCACAGTGTGTTAGATGGTCTTTTATCAACAGTAACCTG	2462
QY	377	ProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProProLeuGln	396	QY	737	LysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPhe	756
Db	1383	CCTGCAAAATCTCTTACAGAGATATCTACAGCTTTGACCGTCTTCCACCACCACTCCAG	1442	Db	2463	AAATTTTCAGAGGCACCATCATGTCTGATTATTTCAGATGCTCGATTTTGGAAAAAGACTTT	2522
QY	397	ProProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeu	416	QY	757	LysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAsp	776
Db	1443	CCTCCTCTGTGAACCTCACTGACCACCGAGAAACAGATTCCACTCTTTTACCATTTCAGTCTC	1502	Db	2523	AAACTATTTAAAAAAATTTTCTCTCTCTGGAATTTAAATATAACAGATTTTACTTGAAGAC	2582
				QY	777	ThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyr	796

QY	239	GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer	258	QY	594	CysTyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrVal	613
DB	1112	CGAGAAACAATAGAAATCTGGAACAGATTATTTCTGTGATGTTTGGCCAGGAAAGAAAGC	1171	DB	2192	FGTTACTTAGACTCAACCTTATTTCTGCTTATTTGCTTTTAGTCTCTGTTCTGACACTGTG	2251
QY	259	LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe	278	QY	614	LeuLeuArgProLysGlyLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeu	633
DB	1172	TTAGGATATTTTGGTGTGACATGGATACCCCTATTGGCACTGGGATGGAAGATTT	1231	DB	2252	TTACTTTAGACCCCAAGAAAGAACGATGTAGAAATATTTATGTGAAACCCCAAGAGCTACTG	2311
QY	279	AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn	297	QY	634	ArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMet	653
DB	1232	GATGAGTGACCTTTGTAGTTTGGGTGTGTGAAGTACAAATCTATTGACATCAAT	1291	DB	2312	AGGACAGAAATGTTTAACTCTCTGAGATATATGATATATGTTGTGTCACAAAAATATG	2371
QY	298	AspIleIlePro-----GluSerValThrGlnGluArgArgProProLysLeuAla	314	QY	654	LysLeuArgLysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLys	673
DB	1292	GATATCATCCAGCTTTATCAGAGAGTGTGCGCAGAAAGAGGCGCTCCCAAACTTGGC	1351	DB	2372	AAACTGAGGAAATATCTTGAAGAGTGGAGGCTGCATCAGGATTTTACCTCTGAAGAAAA	2431
QY	315	PheMetSerArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThr	334	QY	674	AspProGluGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeu	693
DB	1352	TTTATGTCGAAGAGTGTGGGACAAAGGTTTCATCCAGTCATAATAAACAAGGCTACA	1411	DB	2432	GATCTCTGAGGAATCTTTGAATATTTCTGTTTCATCATATTTTAAGGGTAGAACCTTTGCTA	2491
QY	335	GlySerThrSerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGly	353	QY	694	LysIleArgSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGlu	713
DB	1412	GGATCTACCTCAGACCTCGAATAGAAACACAGATCTGAATTTATTTATACCTTAAATGGG	1471	DB	2492	AAATTAAGATCAGCAGGTCAAAAGGTACAAGATTGTTACTTCTATCAAAATTTTATCGAA	2551
QY	354	SerSerValAspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGluVal	373	QY	714	LysAsnGluLysValGlyValProThrIleGlnGlnLeuLeuGluTrpSerPheIleAsn	733
DB	1472	TCTTCTGTTGACTCACACCAACCAATCCAAATCAAAAAATACATGTTGATGAAGTT	1531	DB	2552	AAAAATGAGAAAGTTGGCGTCCCACAATTCAGCAGTTGTAGAAATGGCTTTTATCAAC	2611
QY	374	AlaGluAspProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerPro	393	QY	734	SerAsnLeuLysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGly	753
DB	1532	GCAGAAGACCCCTGCAAAATCTTTACAGAGATATCTACAGACTTTGACCGTCTTCCACCA	1591	DB	2612	AGTAACCTGAAATTTTCAGAGGACCATCATGTCGATTATTCAGATGCCTCGATTGGGA	2671
QY	394	ProLeuGlnProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuPro	413	QY	754	LysAspPheLysLeuPheLysIlePheProSerLeuGluLeuAsnIleThrAspLeu	773
DB	1592	CCACTCCAGCCTCCTCTGTGAACCTCACTGACCACCGAGAACAGATTCCACTCTTTACCA	1651	DB	2672	AAAGACTTTAAACTATTTAAAAAATTTTCTCTCTGGAATTAATAATAACAGATTTA	2731
QY	414	PheSerLeuThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeu	433	QY	774	LeuGluAspThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArg	793
DB	1652	TTTCAGTCTCACCAAGATGCCAATACCAATGGAAGTATTGGCCACAGTCCACTTTCTCTG	1711	DB	2732	CTTTGAAGACACTCCACAGACAGTCCCGATATGTGGAGGGCTTGCAATGTATGAGTGTAGA	2791
QY	434	SerAlaGlnSerValMetGluGluLeuAsnThrAlaProValGlnGluSerProLeu	453	QY	794	GluCysTyrAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCys	813
DB	1712	TCAGCCAGCTCTGTAATGGAAGAGCTAAACACTGCACCCGCTCCAGAGAGTCCACCCCTG	1771	DB	2792	GAAATGCTACGACCATCCGACATCTCAGCTGGAAAAATCAAGCAGATTTTGTAAACCTGC	2851
QY	454	AlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGlu	473	QY	814	AsnThrGlnValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeu	833
DB	1772	GCCATGCCCTCTGGGAACTCACATGCTCTAGAAAGTGGGCTCATTTGGCTGAAGTTAAGGAG	1831	DB	2852	AACTCAAGTCCACCTTTCATCCGAGAGGCTGAATCATTAATATAATACCCAGTGTCACTT	2911
QY	474	AsnProProPheTyrGlyValIleArgTTrpIleGlyGlnProProGlyLeuAsnGluVal	493	QY	834	ProLysAspLeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGlu	853
DB	1832	AACCTCTCTTTATGGGTAAATCCGTGTGATCCGTCAGCCACGACGATGAATGAAGTG	1891	DB	2912	CCCAAGACTTACCCGACTGGGACTGGAGACACGGCTGCATCCCTTCCAGAAATATGGAG	2971
QY	494	LeuAlaGlyLeuGluLeuAspGlyCysAlaGlyCysThrAspGlyThrPheArgGly	513	QY	854	LeuPheAlaValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLys	873
DB	1892	CTCGCTGGACTGGAATGGAAGATGATGTGCGAGGCTGTACGATGGAACCTTCAGAGGC	1951	DB	2972	TTATTTGCTGCTCTCTGCTAGAAAAAACAAGCCACTATGTTGCTTTTGTGAAGTAGTGGGAAG	3031
QY	514	ThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPheValLysLysSerCysArg	533	QY	874	AspAspSerAlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGly	893
DB	1952	ACTCGGTATTTCACTGTGCCCTGGAAGAGCGCTGTTTGTGAAACTCTGAAGAGCTGCAGG	2011	DB	3032	GAGATTCGCCCTGGCTCTTTCTTTGACAGCATGCCCATCGGGATGGTGGTCAAGATGGC	3091
QY	534	ProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSer	553	QY	894	PheAsnIleProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeu	913
DB	2012	CCTGACTCTAGGTTTGATCATTTGACGCGGTTTCCAAATCAGATTGAGCGCTGTAACTCT	2071	DB	3092	TTCAACATTTCTCAGTCAACCCCATGCCAGAGAGTAGAGAGTACTTTGAAGATGTCTCTG	3151
QY	554	LeuAlaPheGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGlu	573	QY	914	GluAspLeuHisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCys	933
DB	2072	TTAGCATTTGAGGCTTACTTAAGTGAAGTAGTAGAGAAATACTCCACCAAAAAATGGAA	2131	DB	3152	GAAGACCTGCTATCTTGGACTCCAGAGAGAAATCCAGGCTGTGCACGAGACTGCTTTGT	3211
QY	574	LysGluGlyLeuGluIleMetIleGlyLysLysLysGlyIleGlnGlyHisTyrAsnSer	593	QY	934	AspAlaTyrMetCysMetTyrGlnSerProThrMetSerLeuTyrLys	949
DB	2132	AAAGAAGGCTTGGAGATATGATTGGGAAGAAAGAGGCATCCAGGGGTCTATTCAATCT	2191	DB	3212	GATGCATATATGTCATGTACACGAGGTCCAAATGAGTTTGTACAAA	3259

ADRI14488
ID ADRI14488 standard; DNA; 5371 BP.
XX
AC ADRI14488;
XX
DT 21-OCT-2004 (first entry)
XX
DE Human NF-kappaB pathway-associated gene SeqID489.
XX
KW NF-kappaB pathway; antiinflammatory; cytostatic; hepatotropic; virucide;
KW antiarthritic; antirheumatic; gastrointestinal-gen; antiaesthetic;
KW antiarteriosclerotic; immunomodulator; cerebroprotective; vasotropic;
KW immunosuppressive; vulnery; gene therapy; immune disorder;
KW inflammatory disorder; NF-kappaB regulation; cancer; aberrant apoptosis;
KW hepatic disorder; Hodgkin's lymphoma; haematopoietic tumour;
KW hyper-IgM syndrome; hypohidrotic ectodermal dysplasia;
KW X-linked anhidrotic ectodermal dysplasia; immunodeficiency;
KW viral infection; HIV-1; HTLV-1; hepatitis B; hepatitis C; EBV; influenza;
KW viral replication; host cell survival; evasion of immune response;
KW rheumatoid arthritis; inflammatory bowel disease; colitis; asthma;
KW atherosclerosis; cachexia; euthyroid sick syndrome; stroke; EAE;
KW autoimmune disorder; hyper immune activity;
KW aberrant acute phase response; hypercongenital condition; birth defect;
KW necrotic lesion; wound; organ transplant rejection;
KW aberrant signal transduction; proliferating disorder; cancer;
KW HIV propagation; gene; ds; human.
XX
OS Homo sapiens.
XX
PN WO2004065577-A2.
XX
PD 05-AUG-2004.
XX
PF 13-JAN-2004; 2004WO-US000798.
PR 14-JAN-2003; 2003US-0440068P.
PR 12-MAY-2003; 2003US-0469757P.
XX
(BRIM) BRISTOL-MYERS SQUIBB CO.
PA
PI Nadler SG, Neubauer MG, Feder JN, Carman J;
XX
XX WPI; 2004-562168/54.
DR P-PSDB; ADRI14489.
XX
XX New isolated polynucleotides and polypeptides associated with NF-kappaB
PT pathway, useful for diagnosing, treating, or preventing disorders or
PT diseases associated with NF-kappaB pathway.
XX
PS Claim 1; SEQ ID NO 489; 237pp; English.
XX
XX This invention relates to the novel association of protein sequences (and
CC the genes which encode them) to the NF-kappaB pathway. The invention may
CC be useful for the production of compounds with an antiinflammatory,
CC cytostatic, hepatotropic, virucide, antiaesthetic, antirheumatic,
CC gastrointestinal-gen, antiaesthetic, antiatherosclerotic,
CC immunomodulator, cerebroprotective, vasotropic, immunosuppressive or
CC vulnery activity or for gene therapy. The proteins and nucleotides are
CC useful for diagnosing, preventing, treating, or ameliorating conditions
CC or diseases associated with the NF-kappaB pathway. The condition is an
CC immune disorder, an inflammatory disorder, an inflammatory disorder
CC related to aberrant NF-kappaB regulation, cancer, aberrant apoptosis,
CC hepatic disorders, Hodgkin's lymphomas, haematopoietic tumours, hyper-IgM
CC syndromes, hypohidrotic ectodermal dysplasia, X-linked anhidrotic
CC ectodermal dysplasia, immunodeficiency, viral infections, HIV-1, HTLV-1,
CC hepatitis B, hepatitis C, EBV, influenza, viral replication, host cell
CC survival, evasion of immune responses, rheumatoid arthritis, inflammatory
CC bowel disease, colitis, asthma, atherosclerosis, cachexia, euthyroid sick
CC syndrome, stroke, EAE, autoimmune disorders, disorders related to hyper
CC immune activity, disorders related to aberrant acute phase responses,
CC hypercongenital conditions, birth defects, necrotic lesions, wounds,
CC organ transplant rejection, conditions related to organ transplant
CC rejection, disorders related to aberrant signal transduction,
CC

CC proliferating disorders, cancers and HIV propagation in cells infected
CC with other viruses. The present sequence is that of a human gene which is
CC subject to the novel association with the NF-kappaB pathway of the
CC invention. Note: This sequence does not appear in the specification but
CC was obtained by the indexer from Genbank.
XX
SQ Sequence 5371 BP; 1600 A; 950 C; 1193 G; 1638 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 0 Length: 5371
Score: 4971.50 Matches: 948
Percent Similarity: 99.16% Conservative: 0
Best Local Similarity: 99.16% Mismatches: 1
Query Match: 98.76% Indels: 7
DB: 13 Gaps: 5
US-09-671-687A-3 (1-949) x ADRI14488 (1-5371)
QY 1 MetSerSerGlyLeuTrpSerGlnGluValThrSerProTyrTrpGluGluArgile 20
DB 392 ATGAGTTTCAGGCTTATGGAGCCCAAGAAAAAGTCACTTCCACCTACTGGGAAGCGGATT 451
QY 21 PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLys 40
DB 452 TTTTACTTGCTTCTTCAGAAATGCGAGCTTACAGACAAACAAACAAAGACTCTTTAAA 511
QY 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgile 60
DB 512 GTACCGAAGGGAAGTATAGACAGATATATTCAAGATCGTTCTGTGGGCACTCAAGGATT 571
QY 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla 80
DB 572 CCTTCTGCAAAAGGCAAGAAAAATCAGATTGGATTAAAAATTTCTAGACCAACCTCATGCA 631
QY 81 ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu 99
DB 632 GTTCTCTTTGTTGATGAAAGGATGTTGTAGAGATAAATGAAAGTTCACAGAGTTACTT 691
QY 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
DB 692 TTGGCAATTACCAATTGTGGAGAGAGGTTTCAAGCTGTTTAAAAACAGACAACTAAGT 751
QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuAArgSerGlyGluGlu 139
DB 752 AAAGGCTTCCAAATAGACGTGGGCTGTCTGTGAAAGTACAGCTGAGATCTGGGGAGAA 811
QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
DB 812 AAATTTCTGGAGTTGTACGCTTTCAGAGGACCCCTGTTCAGAGAGGACAGTCTCCGGA 871
QY 160 IlePhePheGlyValGluLeuLeuGluGluValArgGlyGlnGlyPheThrAspGlyVal 179
DB 872 ATATTCTTTGGAGTTGAATTGCTTGGAAAGAGGTCGTGTTCAAGGTTTTCATGCGGGTG 931
QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
DB 932 TACCAAGGAAACACAGCTTTTTCAGTGTGATGAAGATTGTGGCGGTGTTGTTGCATTCGAC 991
QY 199 LysLeuGluLeuLeuGluAspAspThrAlaLeuGlnSerAspTyrAlaGlyProGly 218
DB 992 AAGCTAGAACTCATAGAAGATGATGACACTTGGCATTTGAAAGATGATTACGCAAGGCTCTGGG 1051
QY 219 AspThrMetGlnValGluLeuProLeuGluIleAsnSerArgValSerLeuLysGly 238
DB 1052 GACCAATGCAAGTTCGAACCTCTCTCTTGGAAATAAATCCAGAGTTTCTTTGAAGGTT 1111
QY 239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
DB 1112 GGAGAAACAATAGAACTCGAAACAGATTATATTCTGTGATGTTTTCAGGAGAAAGAACG 1171
QY 259 LeuGlyTyrPheValGlyValAspMetAspAsnProLysGlyAsnTrpAspGlyArgPhe 278
DB 1172 TTAGGATATTTTGTGTGGGACATGATAACCTTATTTGGCAACTGGGATGGAAGATT 1231

```

QY 279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
DB 1232 GATGGAGTCAGCTTTGTAGTTTGGCTGTGTTGAAAGTACAAATTCATTATTCACATCAAT 1291
QY 298 AspIleIlePro-----GluSerValThrGlnGluArgArgProProLysLeuAla 314
DB 1292 GATATCATCCAGCTTTATCAGAGAGTGTGACGAGAAAGGAGGCTCCCAAACTTGCC 1351
QY 315 PheMetSerArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThr 334
DB 1352 TTTATGTCAAGAGGTGTGGGACAAAGGTTCATCCAGTCATATTAACCAAGGGCTACA 1411
QY 335 GlySerThrSerAspProGlyAsnArg---ArgSerGluLeuPheTyThrLeuAsnGly 353
DB 1412 GGATCTACCTCAGACCCTCGAAATAGAAACAGATCTGAATATTTTATACCTTAATGGG 1471
QY 354 SerSerValAspSerGlnProGlnSerLysSerLysAsnThrTrpTyIleAspGluVal 373
DB 1472 TCTTCTGTGTGACTCACACCAACCAATCCAAATCAAAAATACATGGTACATTGATGAAGTT 1531
QY 374 AlaGluAspProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerPro 393
DB 1532 GCAGAAGACCTCGCAAAATCTCTTACAGAGATATCTACAGACTTTGACCGTTCTTCACCA 1591
QY 394 ProLeuGlnProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuPro 413
DB 1592 CCACCTCCAGCCTCCTCTGTGAACCTCACTGACCACCGAGAAACAGATTCACACTCTTTACCA 1651
QY 414 PheSerLeuThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeu 433
DB 1652 TTCAGTCTCACCAAGATGCCCAATACCAATGGAAGTATTTGGCCACACAGTCCACTTCTCTG 1711
QY 434 SerAlaGlnSerValMetGluGluLeuAsnThrAlaProValGlnGluSerProLeu 453
DB 1712 TCAGCCAGTCTGTAATGGAAGAGCTAAACACTGCACCCGCTCCAGAGAGTCCACCCCTTG 1771
QY 454 AlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGlu 473
DB 1772 GCCATGCTCTCCGGAACTCACATGCTTAGAAGTGGCTCATTTGGCTGAAGTTAAGGAG 1831
QY 474 AsnProProPheTyGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluVal 493
DB 1832 AACCCCTCCTTTCTATGGGGTAAATCCGTGTGATCGGTGAGCCACAGGACTGAATGAAGTG 1891
QY 494 LeuAlaGlyLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGly 513
DB 1892 CTCGCTGGGACTGGAATGGAAGATGAGTGTGACGGCTGTACGGATGGAACCTTCAGAGGC 1951
QY 514 ThrArgTyThrPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArg 533
DB 1952 ACTCGGTATTTACCTGTGCGCTGAAGAGGCGCTGTTGTGAAACTGAAGAGCTGCAGG 2011
QY 534 ProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSer 553
DB 2012 CCTGACTCTAGGTTTGCATCATTTGACGCGGTTTCCAATCAGATTGAGCGCTGTAACCTCT 2071
QY 554 LeuAlaPheGlyGlyTyThrLeuSerGluValValGluGluAsnThrProProLysMetGlu 573
DB 2072 TTAGCATTTGGAGGCTACTTAAGTAGAAGTAGTAGAAGAAATAACTCCACCAAAAATGGAA 2131
QY 574 LysGluGlyLeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyAsnSer 593
DB 2132 AAAGAAGGCTTGGAGATATGATTTGGGAAGAAAGAGGCATCCAGGGTCAATTCAATTCCT 2191
QY 594 CysTyThrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrVal 613
DB 2192 TGTACTTAGACTCAACCTTATTCGTCTATTGCTTTTATGTTCTGTTCTGGACACTGTG 2251
QY 614 LeuLeuArgProLysGluLysAsnAspValGluTyTyThrSerGluThrGlnGluLeuLeu 633
DB 2252 TTACTTAGACCCCAAGAAAGACGATGTAGATATATTATAGTGAACCCCAAGAGCTACTG 2311

```

```

QY 634 ArgThrGluIleValAsnProLeuArgIleTyGlyTyValCysAlaThrLysIleMet 653
DB 2312 AGGACAGAAATTTGTTAATCTCTCGAGATATATGGAATATGTTGTGTGTCACAAAAATATG 2371
QY 654 LysLeuArgLysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLys 673
DB 2372 AAACCTGAGGAAATATCTTGAAAGGTGGAGCTGCATCAGATTTACCTCTGAAGAAAAA 2431
QY 674 AspProGluGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeu 693
DB 2432 GATCCTCGGAATTTCTTGAATATCTTGTTCATCATATTTTAAAGGTAGAACCTTTGCTA 2491
QY 694 LysIleArgSerAlaGlyGlnLysValGlnAspCysTyThrPheTyGlnIlePheMetGlu 713
DB 2492 AAAATTAAGATCAGCAGGTCAAAAGGTACAAGATTTGTACTTCTATCAAAATTTTATGGAA 2551
QY 714 LysAsnGluLysValGlyValProThrIleGlnGlnLeuLeuGluTrpSerPheIleAsn 733
DB 2552 AAAAATGAGAAAGTTGGCGTTCACCAATTCACAGATTGTAGAAATGCTCTTTTATCAAC 2611
QY 734 SerAsnLeuLysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGly 753
DB 2612 AGTAACCTGAAATTTGCAGAGGCACCATCATGTCGTGATTTATTCAGATGCGCTCGATTGGA 2671
QY 754 LysAspPheLysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeu 773
DB 2672 AAAGACTTTAAACTATTTAAAAAAATTTTTCCTTCTCTGGAATTTAAATATAACAGATTTA 2731
QY 774 LeuGluAspThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyThrCysArg 793
DB 2732 CTTTGAAGACACTCCACAGACAGTCCGGATATGTGGAGGGCTTGCAATGTATGATGATAGA 2791
QY 794 GluCysTyThrAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCys 813
DB 2792 GAATGCTACAGCAGTCCGGACATCTCAGCTGGAAAAATCAAGCAGATTTGTTGTAACCTGC 2851
QY 814 AsnThrGlnValHisLeuHisProLysArgLeuAsnHisLysTyThrAsnProValSerLeu 833
DB 2852 AACACTCAAGTCCACCTTCATCCGAAGAGGCTGAATCATATAATAACCCAGTGTCACTT 2911
QY 834 ProLysAspLeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGlu 853
DB 2912 CCAAAAGACTTACCCGACTGGGACTGGAGACACGGCTGCATCCCTTCCCAAGATATGGAG 2971
QY 854 LeuPheAlaValLeuCysIleGluThrSerHisTyThrValAlaPheValLysTyGlyLys 873
DB 2972 TTTATTTGCTGTTCTTGATAGAAACAGCCACTATGTTGCTTTTGTGAAGATATGGGAAG 3031
QY 874 AspAspSerAlaTrpLeuPheAspSerMetAlaAspArgAspGlyGlyGlnAsnGly 893
DB 3032 GACGATTTGCTGGCTCTTCTTTGACAGCATGCGCGATCGGGATGTTGTCAGAATGGC 3091
QY 894 PheAsnIleProGlnValThrProCysProGluValGlyGluTyThrLeuLysMetSerLeu 913
DB 3092 TTTCAACATTTCTCAAGTCAACCCATGCCAGAGAGTAGAGAGTACTTTGAAGATGTCTCTG 3151
QY 914 GluAspLeuHisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCys 933
DB 3152 GAAGACTGTGATTTCTTTGGACTCCAGGAATCAAGGCTGTGACCAAGACTGCTCTTGT 3211
QY 934 AspAlaTyThrMetCysMetTyThrGlnSerProThrMetSerLeuTyLys 949
DB 3212 GATGCATATATGTGATGTACAGAGTCCAAACAATGAGTTTGTACAAA 3259

```

RESULT 6

ADQ95919

ID ADQ95919 standard; cdna; 3311 BP.

XX AC ADQ95919;

XX DT 07-OCT-2004 (first entry)

XX DE T cell activation associated cdna #49.

XX ss; gene; anti-allergic; antiarthritic; antiasthmatic; antidiabetic;
 KW anti-HIV; antimicrobial; antirheumatic; immunosuppressive;
 KW neuroprotective; gene therapy; T cell activation; diagnosis;
 KW autoimmune disease; rheumatoid arthritis; asthma; multiple sclerosis;
 KW diabetes; allergic disease; infectious disease; AIDS; chronic rejection;
 KW organ; bone-marrow transplant.
 XX Homo sapiens.
 OS
 XX
 XX Key Location/Qualifiers
 FH 243..3113
 FT CDS /*tag= a
 FT
 XX
 PN WO2004058805-A2.
 XX
 XX 15-JUL-2004.
 XX
 XX 25-DEC-2003; 2003WO-JP016715.
 XX
 XX 26-DEC-2003; 2002JP-00376365.
 PR 27-DEC-2003; 2002US-0436473P.
 PR 25-APR-2003; 2003JP-00122113.
 PR 28-APR-2003; 2003US-0465792P.
 PR 21-OCT-2003; 2003JP-00360559.
 PR 22-OCT-2003; 2003US-0512846P.
 XX
 XX (ASAH-) ASahi Kasei Pharma Corp.
 PA
 XX Matsuda A, Yoneta S;
 XX WPI; 2004-S93134/57.
 DR P-PSDB; ADQ95920.
 XX
 XX New purified protein involved in T cell activation, useful for
 PT diagnosing, preventing and/or treating acquired immunodeficiency
 PT syndrome, autoimmune (e.g. rheumatoid arthritis, and diabetes), allergic
 PT and infectious diseases.
 PT
 XX Claim 4; SEQ ID NO 97; 2828pp; English.
 PS
 PS The invention relates to purified proteins and genes encoding them, that
 CC are involved in T cell activation (I) and has an amino acid deletion,
 CC substitution or addition in the amino acid sequences. The methods and
 CC compositions of the present invention are useful for the diagnosis,
 CC prevention and/or treatment of autoimmune disease (rheumatoid arthritis,
 CC asthma, multiple sclerosis and diabetes), allergic disease, infectious
 CC disease, AIDS, and acute or chronic rejection at organ transplant or bone
 CC -marrow transplant. This sequence corresponds to a cDNA for a protein
 CC involved in T cell activation.
 XX
 SQ Sequence 3311 BP; 983 A; 650 C; 773 G; 905 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 0 Length: 3311
 Score: 4968.50 Matches: 947
 Percent Similarity: 99.16% Conservative: 1
 Best Local Similarity: 99.06% Mismatches: 1
 Query Match: 98.70% Indels: 7
 DB: 12 Gaps: 5
 US-09-671-687A-3 (1-949) x ADQ95919 (1-3311)
 QY 1 MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluGluArgIle 20
 DB 243 ATGAGTTTCAGGCTATGAGCCAGAAAGATCACTTCACTCCCTACTGGGAAGCGGATT 302
 QY 21 PheTyrLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLeuLys 40
 DB 303 TTTTACTTCTCTTCAAGATTCAGCGTTACAGACAAACAAACAAAGCTCTTTAA 362
 QY 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60

363 GTACCGAAGGAAAGTAGGACAGTATATTCAAGATCGTTCTGTGGGCGCATTTCAAGGATT 422
 QY 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla 80
 DB 423 CCTTCTGCAAAAGGCAAGAAATTCAGATTGGATTAAAAATTTCTAGAGCAACCTCATGCA 482
 QY 81 ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu 99
 DB 483 GTTCTCTTTGTTGATGAAAAGGATGTTGTAGAGATAAATGAAAAGTTTACAGAGTTACTT 542
 QY 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
 DB 543 TTGGCAATTACCAATTGTGAGGAGAGGTTCCAGCTCTGTTAAAAACAGAAAACAGACTAAGT 602
 QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
 DB 603 AAAGGCTCCAAATAGACGTGGGCTGCTCTGTGAAAGTACAGCTCAGATCTGGGGAGAA 662
 QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
 DB 663 AAATTTCTGGAGTTGTACGCTTTCAGAGGAGCCCTGTTAGCAGAGAGACAGTCTCCGGA 722
 QY 160 IlePhePheGlyValGluLeuLeuGluGluGlyArgGlyGlnGlyPheThrAspGlyVal 179
 DB 723 ATATTCTTTGGAGTTGAAATTCCTGGAAGAGGTCGTGCTCAAGGTTTCTACATGACGGGTG 782
 QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
 DB 783 TACCAAGGGAACACAGCTTTTTCAGTGTGATGAAGATTGTGGCGTGTGTTGTCATTGGAC 842
 QY 199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
 DB 843 AAGCTAGAACTCATAGAAGATGATGACACTGCTTGGAAGTGTATTACGAGGCTCTGGG 902
 QY 219 AspThrMetGlnValGluLeuProLeuGluIleAsnSerArgValSerLeuLysGly 238
 DB 903 GACAAATGAGGTGCAAGATTCCTCTTGTGAAATAAACTCCAGAGTTCTTTGAAGTT 962
 QY 239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
 DB 963 GGAGAAACAAATAGAACTCTGGAACAGATTATATCTGTGATGTTTGTGCCAGGAAAGAAAGC 1022
 QY 259 LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe 278
 DB 1023 TTAGGATATTTCTGTTGGTGGACATGGATAACCTATTGGCAACTGGGATGGGAAGATT 1082
 QY 279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
 DB 1083 GATGGAGTGCAGCTTTGTTAGTTTGTGCGTGTGTTGAAAGTACAAATCTATTGCACATCAAT 1142
 QY 298 AspIleIlePro-----GluSerValThrGlnGluArgArgProProLysLeuAla 314
 DB 1143 GATATATCCAGCTTTATTCAGAGAGTGTGACGAGGAAGAGGAGCGCTCCCAAACTTGC 1202
 QY 315 PheMetSerArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThr 334
 DB 1203 TTTATGTCAAGAGGTGTGGGACAAAGGTTTCATCCAGTCATAATAAACCAAGAGCTACA 1262
 QY 335 GlySerThrSerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrIleAsnGly 353
 DB 1263 GGATCTACCTCAGACCCCTGGAAATAGAAACAGATCTCGAATTTATTTATACCTTTAAATGG 1322
 QY 354 SerSerValAspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGluVal 373
 DB 1323 TCTTCTGTTGACTACAAACCAATCCAAATCAAAATAATACATGGTACATTGATGAAGTT 1382
 QY 374 AlaGluAspProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerPro 393
 DB 1383 GCAGAGAGACCTGCATAATCTCTTACAGAGATATCTACAGACTTTGACCGTTCTTACCA 1442
 QY 394 ProLeuGlnProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuPro 413
 DB 1443 CCACTCCAGGCTCTCTCTGTGAACTCACTGACCAACCGAGAACAGATTCCACTCTTTTACCA 1502

```
QY 414 PheSerLeuThrIysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeu 433
Db 1503 TTCAAGTCTCACCAGATGCCCAATACCAATGGAAGTATTGGCCACAGTCCACTTCTCTG 1562
QY 434 SerAlaGlnSerValMetGluGluLeuLeuAsnThrAlaProValGlnGlnSerProLeu 453
Db 1563 TCAGCCAGTCTGTAATGGAAGAGCTAAACACTGACCCGCTCCAGAGAGTCCACCCCTTG 1622
QY 454 AlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGlu 473
Db 1623 GCCATGCTCTGGGAACCTCACATGCTAGAGTGGGCTCATTTGGCTGAAGTTAAGGAG 1682
QY 474 AsnProProPheThrGlyValIleAaGTTpIleGlyGlnProProGlyLeuAsnGluVal 493
Db 1683 AACCCCTCCTTTCTATGGGGTAAATCCGTTGGATCGGTGAGCCACAGGACTGAATGAAGTG 1742
QY 494 LeuAlaGlyLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheAaGly 513
Db 1743 CTCGCTGGGACTGGAATCTGGAAGATGAGTGTGCAAGGCTGTACGGATGGAACCTTCAGAGGC 1802
QY 514 ThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArg 533
Db 1803 ACTCGGTATTTACCTGTGCTTGAAGAGGCGCTGTTGTGAACCTGAAGAGCTGCAGG 1862
QY 534 ProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSer 553
Db 1863 CCTGACTCTAGGTTTGCATCATTGCAAGCGGTTTCCAAATCAGATTGAGCGCTGTAACCTCT 1922
QY 554 LeuAlaPheGlyClyTyrLeuSerGluValValGluGluAsnThrProProLysMetGlu 573
Db 1923 TTAGCATTTGGAGGCTACTTAAGTAGAGTAGTAGAAGAAATACTCCACCACAAATAATGGA 1982
QY 574 LysGluGlyLeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSer 593
Db 1983 AAGAAGGCTTGGAGATAATGATTGGGAAGAGGCAATCCAGGCTCAATTACAATCT 2042
QY 594 CysTyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrVal 613
Db 2043 TGTACTTAGACTCAACCTTATCTGCTTATTTGCTTTTAGTTCTGTTCTGGACACTGTG 2102
QY 614 LeuLeuArgProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeu 633
Db 2103 TTACTTTAGACCCCAAGAAAGACGATGTAGAAATATTATAGTGAAACCCCAAGAGCTACTG 2162
QY 634 ArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMet 653
Db 2163 AGGACAGAAATTTGTTAATCTCTGAGAAATATATGGATATATGTGTGCCACAAATAATTATG 2222
QY 654 LysLeuArgLysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLys 673
Db 2223 AACTGAAGAAATACTTGAAGAGTGGAGCTGCATCAGATTACTCTGAAGAAAA 2282
QY 674 AspProGluGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeu 693
Db 2283 GATCCTGAGGAATCTTGAATATCTGTTTCATCATATTTAAGGGTAGAACCTTTGCTA 2342
QY 694 LysIleArgSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGlu 713
Db 2343 AAAATAAGATCAGCAGGTCAAAAGGTACAAAGATTGTTACTTCTATCAAAATTTTATGGAA 2402
QY 714 LysAsnGluLysValGlyValProThrIleGlnLeuLeuGluTyrPhePheIleAsn 733
Db 2403 AAAAATGAGAAGTTGGCGTTCCACAAATTCAGCAGTTGTAGATGGTCTTTTATCAAC 2462
QY 734 SerAsnLeuLysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGly 753
Db 2463 AGTAACCTGAAATTTGCAGAGGCACCATCATGCTGATTATTTCAGATGCGCTCGATTGGA 2522
QY 754 LysAspPheLysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeu 773
Db 2523 AAAGACTTTAACTATTTAAAAAATTTTCTCTCTGGAATTAATAATAACAGATTTA 2582
```

```
QY 774 LeuGluAspThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArg 793
Db 2583 CTTGAAGACACTCCCAAGACAGTCCCGGATATGTGGAGGGCTTGCATATGATGATGATGA 2642
QY 794 GluCysTyrAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCys 813
Db 2643 GAATGCTACGACGATCCGGACATCTCAGCTGGAAAAATCAAGCAGTTTGTAAACCTGC 2702
QY 814 AsnThrGlnValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeu 833
Db 2703 AACACTCAAGTCCACCTTCATCCGAAGAGGCTGAATCATATAATACCCAGTGTCACTT 2762
QY 834 ProLysAspLeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGlu 853
Db 2763 CCCAAAGACTTACCGACTGGGACTGGAGACAGCGCTGCATCCCTTCCAGAAATATGGAG 2822
QY 854 LeuPheAlaValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLys 873
Db 2823 TTATTTGCTGTTCTCTGCATAGAAACAGCCACTATGTTGCTTTTGTGAAGTATGGGAAG 2882
QY 874 AspAspSerAlaTrpLeuPheAspSerMetAlaAspArgAspGlyGlyGlnAsnGly 893
Db 2883 GACGATTCTGCTGGCTCTTCTTTGACAGCATGCGCGATCGGTGTCAGATGGC 2942
QY 894 PheAsnIleProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeu 913
Db 2943 TTCAACATTTCTCAAGTCAACCCATGCCAGAAAGTAGGAGTACTTTGAAGATGCTCTGT 3002
QY 914 GluAspLeuHisSerLeuAspSerArgIleGlnGlyCysAlaArgLeuLeuCys 933
Db 3003 GAAGAGCTGCTATCTCTGGACTCCAGAGAAATCAAGGCTGTGCACGAGACTGCTTGT 3062
QY 934 AspAlaTyrMetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
Db 3063 GATGCATATATGTGCATGTACAGAGTCCACAAATGATGATTTGTACAAA 3110

RESULT 7
AAA08589
ID AAA08589 standard; DNA; 4527 BP.
XX
AC AAA08589;
XX
DT 19-JUL-2000 (first entry)
XX
DE Human cytoskeleton associated protein 9 (CVSKP-9) coding sequence.
XX
KW Cytoskeleton associated protein; CVSKP-9; cancer; proliferative;
KW autoimmunity; inflammatory; vesicle trafficking; neurological;
KW cardiovascular; cell motility; reproductive; muscle disorder; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 926..3121
FT /*tag= a
FT /product= "CVSKP-9"
XX
XX WO200017355-A2.
XX
XX 30-MAR-2000.
XX
XX 17-SEP-1999; 99WO-US021565.
XX
XX 18-SEP-1998; 98US-0172226P.
XX 27-APR-1999; 99US-0131321P.
XX
XX (INCY-) INCYTE PHARM INC.
XX
XX Lal P, Tang YT, Yue H, Hillman JL, Bandman O, Corley NC;
XX Guegler KJ, Patterson C, Azimzai Y, Baughn MR;
XX WPI; 2000-283582/24.
XX
XX P-PSDB; AAY91954.
```

102 (e) date
3 No
available

XX Human cytoskeleton associated proteins, used to treat cell proliferative,
 PT autoimmune/inflammatory, vesicle trafficking, neurological, cell
 PT motility, reproductive and muscle disorders.

XX Claim 9; Page 107-109; 113pp; English.

XX AAA08581-96 encode human cytoskeleton associated proteins 1 to 16 (CYSKP-
 CC 1 to CYSKP-16) respectively. The sequences can be used to treat and
 CC diagnose cancer and cell proliferative, autoimmune/inflammatory, vesicle
 CC trafficking, neurological, cardiovascular, cell motility, reproductive
 CC and muscle disorders. Pharmaceutical compositions containing CYSKP-2 to
 CC CYSKP-16 can be used to treat or prevent disorders associated with
 CC decreased expression or activity of CYSKP (claimed), for example,
 CC atherosclerosis, cirrhosis, hepatitis, myelofibrosis, psoriasis, cancers,
 CC autoimmune/antiinflammatory disorders such as allergies, anemia, asthma,
 CC acquired immunodeficiency syndrome (AIDS), Crohn's disease, diabetes
 CC mellitus, Goodpasture's syndrome, osteoarthritis, scleroderma and trauma.
 CC CYSKP antagonists can be used to treat or prevent a disorder associated
 CC with increased expression or activity of CYSKP (claimed)

SX Sequence 4527 BP; 1351 A; 823 C; 1005 G; 1348 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0 Length: 4527
 Score: 4942.00 Matches: 944
 Percent Similarity: 99.37% Conservative: 0
 Best Local Similarity: 99.37% Mismatches: 2
 Query Match: 98.17% Indels: 5
 DB: 3 Gaps: 4

US-09-671-687A-3 (1-949) x AAA08589 (1-4527)

QY	4	GlyLeuTrpSerGlnGluLysValThrSerProTyrTriPLeuGluArgIlePheTyrLeu	23
DB	270	GGCTTATGAGGCCAGAAAAGTCATCTCACCTACTGGGAAGCGGATTTTACTTG	329
QY	24	LeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLysValProLys	43
DB	330	CTTCTTCAAGATGCGCGCTTACAGACAAACAAACAAAGCTCTTAAAGTACCAGAA	388
QY	44	GlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIleProSerAla	63
DB	389	GGAAGTATAGGACAGTATATTCAGATCGTCTGTGGGCGATTCAGAGATTCCTCTGCA	448
QY	64	LysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAlaValLeuPhe	83
DB	449	AAAGGCAAGAAAATCAGATTGATTGATTAATAATTTCTAGAGCAACTCATGCGAGTCTCTTT	508
QY	84	ValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeuAlaIle	102
DB	509	GTTGATGAAAGGATGTTGTAGAGATAAATGAAAGTTACAGAGTTACTTTGGCAATT	568
QY	103	ThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSerLysGlyLeu	122
DB	569	ACCAATTGTGAGCAGAGGTTACGCTGTTTAAAAACAGAAACAGACTAAGTAAAGGCTC	628
QY	123	GlnIleAspValCysProValLysValGlnLeuArgSerGlyGluLysPhePro	142
DB	629	CAAAATAGACGTGGGCTGCTCTGTGAAGTACAGCTGAGATCTGGGGAAGAAAATTCCT	688
QY	143	GlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGlyIlePhePhe	162
DB	689	GGAGTTGTACGTTTACAGAGACCCCTTTAGCAGAGAGGACAGTCTCCGGAATATCTTT	748
QY	163	GlyValGluLeuLeuGluGlyArgGlyGlnGlyPheThrAspGlyValTyrGlnGly	182
DB	749	CGAGTTGAATTGCTGGAAGAGGTCTGTGTCAAGGTTTCACTCACGGGGGTATACCAAGGG	808
QY	183	LysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAspLysLeuGlu	201
DB	809	AAACAGCTTTTTCAGTGTGATGAAGATTGTGGCGTGTGTTGTTGCAATTGGACAAGCTAGAA	868
QY	202	LeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGlyAspThrMet	221
DB	869	CTCATAGAAGATGATGACTGTCATTGGAAGTGAATACGCGAGTCTCTGGGACACAATG	928
QY	222	GlnValGluLeuProLeuGluIleAsnSerArgValSerLeuLysGlyGlyGluThr	241
DB	929	CAGGTGCAACTTCTCTCTTGGAAATAAATCCAGAGTTCTTTGAAAGTTGGAGAACACA	988
QY	242	IleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSerLeuGlyTyr	261
DB	989	ATAGAATCTGGAACAGTTATATTCTGTGATGTTTGGCAGGAAAAGAAAGCTTAGGATAT	1048
QY	262	PheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPheAspGlyVal	281
DB	1049	TTTGTGTGTGACATGGATAACCCATTGGCAACTGGGATGGAAGATTGATGGAGTG	1108
QY	282	---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsnAspIleIle	300
DB	1109	CAGCTTTGTAGTTTGGTGTGTTGAAAGTACATTTCTATTGGACATCAATGATATCATC	1168
QY	301	ProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMetSerArgGlyVal	320
DB	1169	CCAGAGAGTGTGACGCGAGGAAAGGAGGCTCCCAAACTTGCTTTATGTCAAGAGGTGT	1228
QY	321	GlyAspLysGlySerSerSerHisAsnLysProLysAlaThrGlySerThrSerAspPro	340
DB	1229	GGGACAAAGGTTTCATCCAGTCATAATAAACCAAGGCTACAGGATCTTACCTCAGACCT	1288
QY	341	GlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerValAspSerGln	359
DB	1289	GGAAATAGAAACAGATCTGAATTTATACCTTTAAATGGGTCTTCTGTGACTCACA	1348
QY	360	ProGlnSerLysSerLysAsnThrTriPtyrIleAspGluValAlaGluAspProAlaLys	379
DB	1349	CCACATCCAAATCAAAAATACATGTTACATTTGATGAAGTTGCAGAGAGCCCTGCAAA	1408
QY	380	SerLeuThrGluLeuSerThrAspPheAspArgSerSerProProLeuGlnProPro	399
DB	1409	TCTCTTACAGAGATATCTACAGACTTTTGACCGTTCTTCCACCACCACTCCAGCCTCCTCT	1468
QY	400	ValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeuThrLysMet	419
DB	1469	GTGAACCTCACTGACCCAGAGAGACAGATTCACCTCTTTTACCATTTCAGTCTCACCAGATG	1528
QY	420	ProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGlnSerValMet	439
DB	1529	CCCAATACCAATGGAAGTATTTGCCACAGTCCACTTTCTGTGTCAGCCAGCTCTGTAATG	1588
QY	440	GluGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetProProGlyAsn	459
DB	1589	GAAGAGCTAAACACTGCACCCGTCCAAGAGAGTCCACCTTTGGCCATGCCTCTCTGGGAAC	1648
QY	460	SerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProProPheTyrGly	479
DB	1649	TCATGCTGTAGAAAGTGGGCTCATTTGGCTGAAGTTAAGGAGAACCCCTCTCTTTCTATGGG	1708
QY	480	ValIleArgTriPleGlyGlnProProGlyLeuAsnGluValLeuAlaGlyLeuGluLeu	499
DB	1709	GTAAATCGTTGGATCGGTGCGTACGCCAGGACCTGAATGAAGTGCTCGCTGGAGTGAATG	1768
QY	500	GluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyrPheThrCys	519
DB	1769	GAAGATGAGTGTGCGGCTGTACGGATGGAACCTTCAGAGGCACTCGGTATTTCACTGT	1828
QY	520	AlaLeuLysLysAlaLeuPheValLysLysSerCysArgProAspSerArgPheAla	539
DB	1829	GCCCTGAAGAGGCGCTGTTTGTGAAACTGAAGAGCTGCAGGCTGACTCTAGGTTTGA	1888
QY	540	SerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPheGlyTyr	559
DB	1889	TCATTTGAGCGGCTTCCCAATTCAGATTCAGCGCTGTAACTCTTTAGCATTTGGAGGCTAC	1948
QY	560	LeuSerGluValValGluGluAsnThrProProLysMetGluLysGluGlyLeuGluIle	579

Db 1949 TTAAGTGAAGTAGAGAGAAATACTCCACAAAATGGAAGAGGCTTGGAGATA 2008
QY MetIleGlyLysLysGlyLysGlnGlyHisTyrAsnSerCysTyrLeuAspSerThr 599
Db 2009 ATGATTGGGAAGAAGAAGGCATCCAGGGTCATTACAATCTCTGTACTTAGACTCAACC 2068
QY LeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArgProLysGlu 619
Db 2069 TTATTCCTGCTATTCTGCTTTAGTCTGTCGACACTGTGTACTTAGACCCAAAGAA 2128
QY LysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGluIleValAsn 639
Db 2129 AGAACGATGTAGAAATATTATAGTGAACCCCAAGAGCTACTGAGGACAGAAATGTTAAT 2188
QY ProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArgLysIleLeu 659
Db 2189 CCTCTGAGAAATATAGTATGTGTGTGCCACAAAATATGAACTGAGAAAATACIT 2248
QY GluLysValGluAlaSerGlyPheThrSerGluGluLysAspProGluGluPheLeu 679
Db 2249 GAAAGGTGGAGCTGCATCAGGATTTACCTCTGAAGAAAAGATCCTGAGGAATCTTG 2308
QY AsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArgSerAlaGly 699
Db 2309 AATATTCTGTTTCATCATATTTTAAAGGGTAGAACCTTTGCTAAAAATAAGATCAGCAGGT 2368
QY GlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGluLysValGly 719
Db 2369 CAAAAGGTACAGATTTGTTACTTCTATCAAAATTTTATGAAAAAATGAGAAAGTTGGC 2428
QY ValProThrIleGlnGlnLeuLeuGluTyrPheIleAsnSerAsnLeuLysPheAla 739
Db 2429 GTTCCCACAATTCAGCAGTTGTTAGAAATGCTCTTTATCAACAGTAACCTGAAATTTGCA 2488
QY GluAlaProSerCysIleuIleGlnMetProArgPheGlyLysAspPheLysLeuPhe 759
Db 2489 GAGGCACCATCATGCTCTGATTAATTCAGATCGCTCGATTTTGGAAAAAGACTTTAAACTATTT 2548
QY LysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAspThrProArg 779
Db 2549 AAAAAAATTTTCTCTCTCGAAATTAATAATATACAGATTTACTTGAAGACACTCCCGA 2608
QY GlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyrAspAspPro 799
Db 2609 CAGTGCAGATATGTGGGGCTTGCAATGTATGATGATAGAAATGCTACGACGATCCG 2668
QY AspIleSerAlaGlyLysLysGlnPheCysLysThrCysAsnThrGlnValHisLeu 819
Db 2669 GACATCTCAGCTGCAAAAATCAAGCAGTTTGTGTAACCTGCAACACTCAAGTCCACCTT 2728
QY HisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAspLeuProAsp 839
Db 2729 CATCCGAAGAGGCTGAATCAATAATAATAACCCAGTGTCACTTCCCAAGACTTACCCGAC 2788
QY TrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAlaValLeuCys 859
Db 2789 TGGACTGGAGACAGCGCTGCATCCCTTGCAGAAATAGAGTATTGTCTGTCTCTGC 2848
QY IleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspSerAlaTrpLeu 879
Db 2849 ATAGAAACAGCCACTATGTTGCTTTGTGAAGTATGGGAAGCAGGATCTGCTGGCTC 2908
QY PhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIleProGlnVal 899
Db 2909 TTCTTTTACACAGCATGCGCATGCGGATGGTGGTGCAGAAATGGCTTCAACATTCCTCAAGTC 2968
QY ThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLeuHisSerLeu 919
Db 2969 ACCCCATGCCAGAGTAGAGAGATGTTGGAAGATGTTCTTGGAGACCTGCAATTCCTTG 3028
QY AspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyrMetCysMet 939

Db 3029 GACTCAGAGAAATCCAGGCTGTGCACGAGACTGCTTTGTGATGATATATGTGCATG 3088
QY 940 TyrGlnSerProThrMetSerLeuTyrLys 949
Db 3089 TACCAGATCCAACAATGAGTTTGTACAAA 3118
RESULT 8
AAH18625
ID AAH18625 standard; cDNA; 2845 BP.
XX
AC AAH18625;
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA sequence SEQ ID NO:18842.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
(HELI-) HELIX RES INST.
XX
XX Ota T, Isoqai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.
XX
PS Claim 8; SEQ ID NO 18842; 2537pp + Sequence Listing; English.
XX
CC The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention
XX
SQ Sequence 2845 BP; 854 A; 557 C; 662 G; 772 T; 0 U; 0 Other;
Alignment Scores: 0 Length: 2845
Pred. No.: 0

Score: 4601.00 Matches: 875
Percent Similarity: 99.32% Conservative: 0
Best Local Similarity: 99.32% Mismatches: 2
Query Match: 91.40% Indels: 4
DB: 4 Gaps: 4

US-09-671-687A-3 (1-949) x AAH18625 (1-2845)

QY	73	LysIleLeuGluGlnProHisAlaValLeuPheValAspGlu---AspValValGluIle	91
Db	2	AAATTTCTAGAGCAACCTCATGTCAGTCTCTTTGTTGATCAAAAGCATGTTGTAGAGATA	61
QY	92	AsnGluLysPheThrCluIleuLeuLeuAlaIleThrAsnCysGluGluArgPheSerLeu	111
Db	62	AATGAAAAGTTACAGAGTTACTTTTGGCAATTACCAATTGTGAGAGAGTTTCAGCCCTG	121
QY	112	PheLysAsnArgAsnArgLeuSerLysGlyLeuGlnIleAspValGlyCysProValLys	131
Db	122	TTTTAAACACAGAACAGACTTAGTAAAGGCTCCAAATAGACGTGGCTGTCTCTGTGAAA	181
QY	132	ValGlnLeuArgSerGlyGluGluLysPheProGlyValValArgPheArgGlyProLeu	151
Db	182	GTACAGCTCAGATCTGGGGAAGAAAATTTCTGGAGTTGTACGCTTCAGAGGACCCCTG	241
QY	152	LeuAlaGluArgThrValSerGlyIlePhePheGlyValGluLeuLeuGluGlyArg	171
Db	242	TTAGCAGAGAGACAGTCTCCGGAATATTCTTTGGAGTTGCAATTGCTGGAAGAGGTCGT	301
QY	172	GlyGlnGlyPheThrAspGlyValTyrglnGlyLysGlnLeuPheGlnCysAspGluAsp	191
Db	302	GGTCAAGGTTTCACTGACCGGGGTACCAAGGGAACACAGCTTTTTCAGTGTGATGAAGAT	361
QY	192	CysGly---PheValAlaLeuAspLysLeuGluLeuIleGluAspAspThrAlaLeu	210
Db	362	TGTGGCGTGTGTTGTCATTGACAGCTAGAACTCATAGAAGATGATGACACTGCATTG	421
QY	211	GluSerAspTyralaGlyProGlyAspThrMetGlnValGluLeuProLeuGluIle	230
Db	422	GAAGTGATTACGCGAGTCTCTGGGACACAAATGACAGTCTCTCTCTTTGGAAATA	481
QY	231	AsnSerArgValSerLeuLysGlyGlyGluThrIleGluSerGlyThrValIlePheCys	250
Db	482	AACTCCAGAGTTCTTTGTAAGGTGGAGAAACAATAAGATCTTGAAACAGTTATATTCTGT	541
QY	251	AspValLeuProGlyLysGluSerLeuGlyTyrrPheValGlyValAspMetAspAsnPro	270
Db	542	GATGTTTTCCAGGAAAGAAAGCTTAGGATATTTTGGTGTGACATGGATACCCCT	601
QY	271	IleGlyAsnTrpAspGlyArgPheAspGlyVal---LeuCysSerPheAlaCysValGlu	289
Db	602	ATTGGCACTGGGATGGAGATTGATGGAGTGCAGCTTTTGTAGTTTTCGGTGTGTGAA	661
QY	290	SerThrIleLeuLeuHisIleAsnAspIleIleProGluSerValThrGlnGluArgArg	309
Db	662	AGTACAAATTTCTATTGACATCAATGATATCATCCAGAGAGTGTGACGCGAGAAAGAGG	721
QY	310	ProProLysLeuAlaPheMetSerArgGlyValGlyAspLysGlySerSerHisAsn	329
Db	722	CCTCCCAAATTTGCCCTTTATGTCAAGAGGTGTGGGGCAAAAGGTTTCACCACTATAAT	781
QY	330	LysProLysAlaThrGlySerThrSerAspProGlyAsnArg---ArgSerGluLeuPhe	348
Db	782	AAACCAAGGCTACAGGATCTACTCAGACCTTGGAAATAGAAACAGATCTGAAATATT	841
QY	349	TyrThrLeuAsnGlySerSerValAspSerGlnProGlnSerLysSerLysAsnThrTrp	368
Db	842	TATACCTTAATGGGCTCTCTGTGACTCACAACCACTCCAAATCCAAATAATACATGG	901
QY	369	TyrIleAspGluValAlaGluAspProAlaLysSerLeuThrGluIleSerThrAspPhe	388
Db	902	TACATTGATGAAGTTGCAGAGACCCCTGCAAAATCTCTTACAGAGATATCTACAGACTTT	961
QY	389	AspArgSerSerProLeuGlnProProValAsnSerLeuThrThrGluAsnArg	408

Db	962	GACCGTTCTTACACACACTCCAGCCTCCTCTCTGTGAACCTCACTGACCCGAGAACAGA	1021
QY	409	PheHisSerLeuProPheSerLeuThrLysMetProAsnThrAsnGlySerIleGlyHis	428
Db	1022	TTCCACTCTTTTACCATTCACTCTCACCAAGATGCCAATACCAATGGAAGTATTGGCCAC	1081
QY	429	SerProLeuSerLeuSerAlaGlnSerValMetGluGluLeuAsnThrAlaProValGln	448
Db	1082	AGTCCACTTTCTCTGTGAGCCAGTCTGTATATGAGAGACTTAAACACTGCACCCCTCCAA	1141
QY	449	GluSerProProLeuAlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeu	468
Db	1142	GAGAGTCCACCTTGGCCATGCTCTCTGGGAACCTACATGCTGTAGAAAGTGGGCTCATGT	1201
QY	469	AlaGluValLysGluAsnProProPheTyrrGlyValIleArgTrpIleGlyGlnProPro	488
Db	1202	GCTGAAGTTAAGGAGAACCTCTCTCTTCTATGGGGTAAATCCGTTGGATCGGTCAGCACA	1261
QY	489	GlyLeuAsnGluValLeuAlaGlyLeuGluLeuGluAspGluCysAlaGlyCysThrAsp	508
Db	1262	GGACTGAATGAAGTCTGCTGAGCTGGAACCTGGAAGATGAGTGTGACAGGCTGACGGAT	1321
QY	509	GlyThrPheArgGlyThrArgTyrrPheThrCysAlaLeuLysLysAlaLeuPheValLys	528
Db	1322	GGAACTTTCAGAGCACTCGGTATTTCACCTGTGCTTGAAGAGGCGCTGTGTGTGAAA	1381
QY	529	LeuLysSerCysArgProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIle	548
Db	1382	CTGAAGAGCTGAGGCTGACTCTAGGTTTGCATCATTCAGCCGCTTCCATCAGATT	1441
QY	549	GluArgCysAsnSerLeuAlaPheGlyTyrrLeuSerGluValValGluGluAsnThr	568
Db	1442	GAGCGCTGTAACCTTTTAGCATTTGGAGGCTACTTAAAGTGAAGTAGTAGAAGAAATACT	1501
QY	569	ProProLysMetGluLysGluGluIleMetIleGlyLysLysLysGlyIleGln	588
Db	1502	CCACCAAAATGGAAGAGAGGCTTGGAGATATGATTTGGGAAGAGAAAGGATCCAG	1561
QY	589	GlyHisTyrrAsnSerCysTyrrLeuAspSerThrIleuPheCysLeuPheAlaPheSer	608
Db	1562	GGTCATTACAAATCTTGTATTCTAGACTCAACCTTATTCTGCTTATTTCGCTTTTAGTTCT	1621
QY	609	ValLeuAspThrValLeuLeuArgProLysGluLysAsnAspValGluTyrrTyrrSerGlu	628
Db	1622	GTTCTGGACACTGTGTACTTTGGACCCAAAGAAAGAACGATGTAGAATATNTAGTGA	1681
QY	629	ThrGlnGluLeuLeuArgThrGluIleValAsnProLeuArgIleTyrrGlyTyrrValCys	648
Db	1682	ACCAAGAGCTACTGAGGACAGAAAATTTGTTAAATCTCTGAGAAATATATGGATATGTGT	1741
QY	649	AlaThrIysIleMetLysLeuArgLysIleLeuGluLysValGluAlaAlaSerGlyPhe	668
Db	1742	GCCACAAAATTTATGAAACTGAGGAAAATACTTGAAGAGGTGGAGGCTGCATCAGGATTT	1801
QY	669	ThrSerGluGluLysAspProGluGluPheLeuAsnIleLeuPheHisIleLeuArg	688
Db	1802	ACCTCTGAAGAAAAGATCTGAGGAATCTTGAAATATCTGTTTCATCATATTTTAAGG	1861
QY	689	ValGluProLeuLeuLysIleArgSerAlaGlyGlnLysValGlnAspCysTyrrPheTyrr	708
Db	1862	GTAGAACCCTTGTGTAATAAATAAGATCAGCAGGTCAAAAGGTACAAGATTGTTACTTAT	1921
QY	709	GlnIlePheMetGluLysAsnGluLysValGlyValProThrIleGlnGlnLeuLeuGlu	728
Db	1922	CAAAATTTTATGAAAAAATGAGAAAGTTGGCGTTCCCAAAATTCAGCAGTGTGTAGAA	1981
QY	729	TrpSerPheIleAsnSerAsnLeuLysPheAlaGluAlaProSerCysLeuIleIleGln	748
Db	1982	TGCTCTTTTATCAACAGTAACTTGAATTTGCAGAGGCAACCATCATGTCTGATTATT	2041
QY	749	MetProArgPheGlyLysAspPheLysLeuPheLysLysIlePheProSerLeuGluLeu	768

QY 401 AsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeuThrLysMetPro 420
DB 572 AACTCTACTGACCACCGAGACAGATTCCACTCTTTTACCAATTCAGTCTCACCAAGATGCC 631
QY 421 AsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGlnSerValMetGlu 440
DB 632 AATACCAATGGAAAGTATTGGCCACAGTCCACTTTCTGTGCACCCAGGCTCTGTAATGGAA 691
QY 441 GluLeuAsnThrAlaProValGlnGlnSerProProLeuAlaMetProProGlyAanSer 460
DB 692 GAGCTAAACACATGACCCCGTCCAAAGAGAGTCCACCCCTTGCCCATGCTCTCTGGNACTCA 751
QY 461 HisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProProPheThrGlyVal 480
DB 752 CATGGTCTAGAAGTGGGCTCATTTGGCTGAAGTTAAGAGAACCTCTCTTCTATGGGGTA 811
QY 481 IleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGlyLeuGluLeuGlu 500
DB 812 ATCCGTTGGATCGGTGAGCCACGAGCTGAATGAAGTGTCTGCTGGACTGGAACTGGAA 871
QY 501 AspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyrPheThrCysAla 520
DB 872 GATGAGTGTGAGGCTGTACGGATGGAACTTCAGAGGCACTCGGTATTTCACCTGTGCC 931
QY 521 LeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSerArgPheAlaSer 540
DB 932 CTGAAGAGCGCGTGTGTGAACTGAAGAGCTGAGAGGCTGACCTAGGTTCATCA 991
QY 541 LeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPheGlyGlyTyrLeu 560
DB 992 TTGCAGCCGCTTTCCAAATCAGATTGAGCGCTGTAACCTTTAGCATTTGGAGGCTACTTA 1051
QY 561 SerGluValValGluGluAsnThrProProLysMetGluLysGluGlyLeuGluIleMet 580
DB 1052 AGTGAAGTAGTAGAAGAAATACTCCACCAAAATGGAAGAGGCTTGGAGATATG 1111
QY 581 IleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeuAspSerThrLeu 600
DB 1112 ATTGGGAAGAAAGGATCCAGGCTCATTAACAATTTCTTTTACTTAGACTCAACCTTA 1171
QY 601 PheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArgProLysGluLys 620
DB 1172 TTCTGCTTATTGCTTTTGTAGTTCTGTCTGGACACTGTGTACTAGACCCAAAGAAAG 1231
QY 621 AsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGluIleValAsnPro 640
DB 1232 AACGATGTAGAATATTATAGTGAACCCAGAGCTACTGAGGACAGAAATTTGTTATCCT 1291
QY 641 LeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArgLysIleLeuGlu 660
DB 1292 CTGAGAATATATGATATGTGTGCCACCAAAATTTATGAACCTGAGGAAATTTGAA 1351
QY 661 LysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGluGluPheLeuAsn 680
DB 1352 AAGGTGGAGGCTGCATCAGGATTTACCTCTGAAGAAAAAGATCCTCAGGAAATTTCTGAAT 1411
QY 681 IleLeuPheHisIleLeuArgValGluProLeuLysIleArgSerAlaGlyGln 700
DB 1412 ATTCTGTTTCATCATATTTTAAAGGGTAGAACCTTTGCTAAAAATAAGATCAGAGGTCAA 1471
QY 701 LysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGluLysValGlyVal 720
DB 1472 AAGGTACAGATTGTACTTCTATCAAAATTTTATGGAAAAAATGAGAAATTTGGCGTT 1531
QY 721 ProThrIleGlnGlnLeuLeuGluTyrPheIleAsnSerAsnLeuLysPheAlaGlu 740
DB 1532 CCCACAATTCAGCAGTGTGTAGAATGGTCTTTTATCAACAGTAACCTGAAATTTGCAGAG 1591
QY 741 AlaProSerCysIleIleGlnMetProArgPheGlyLysAspPheLysLeuPheLys 760
DB 1592 GCACCATCATGCTGTGATTATTCAGATGCCCTCGATTTGGAAAAAGACTTTAAACTATTAA 1651
QY 761 LysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAspThrProArgGln 780

DB 1652 AAAATTTTCTCTCTCGAAATTAACAGATTTTACTTGAAGACACTCCAGACAG 1711
QY 781 CysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyrAspAspProAsp 800
DB 1712 TGCCGGATATGGAGGGCTTGCATATGATGAGTGTAGAGATGCTACGATGATCCGGAC 1771
QY 801 IleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGlnValHisLeuHis 820
DB 1772 ATCTCAGCTGGAATAATCAAGCAGTTTGTAAAACTGCAACACTCAAGTCCACCTTCAT 1831
QY 821 ProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAspLeuProAspTrp 840
DB 1832 CCGAAGAGGCTGAATCATATAATAACCCAGTGTCTCTCCCAAGACTTACCCGACTGG 1891
QY 841 AspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAlaValLeuCysIle 860
DB 1892 GACTGAGACACCGCTGTCATCCCTTGCAGAAATATGAGTATTGTCTGTCTCTGCATA 1951
QY 861 GluThrSerHisTyrValAlaPheValLysTyrGlyLysAspSerAlaTrpLeuPhe 880
DB 1952 GAAACAAGCCACTATGTTGCTTTTGTGAAGTATGGGAAGGACGATTTCTGCTGCTCTTC 2011
QY 881 PheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIleProGlnValThr 900
DB 2012 TTTGACAGCATGGCCGATCGGATGGTGTGTCAGATGGCTTCAACATTTCTCAAGTCACC 2071
QY 901 ProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLeuHisSerLeuAsp 920
DB 2072 CCATGCCCAAGATAGAGAGTACTTGAAGATGTCTCTGGAAGACCTGCAATTTCTTGAC 2131
QY 921 SerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyrMetCysMetTyr 940
DB 2132 TCAGAGAGAAATCCAGGCTGTGCACGAGACTGCTTTGTGATGCATATATGTCATGTAC 2191
QY 941 GlnSerProThrMetSerLeuTyrLys 949
DB 2192 CAGAGTCCCAACATGAGTTGTACAAA 2218
RESULT 10
AA158410
ID AA158410 standard; cDNA; 2523 BP.
XX
AC AA158410;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 613.
XX
KW Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
OS Homo sapiens.
XX
PN W0200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US034263.
XX
PR 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.

XX	(HYSE-) HYSEQ INC.	465	CCACCACCTCCAGCCTCCTCCTGTGAACACTCACTGACCACCGAGAACAGATTCCACTCTTTA	524
XX		413	ProPhSerLeuThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSer	432
PI	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;	525	CCATTGAGTCTCACCAGATGCGCAATACCAATGGAAGTATTGGCCACAGTCCACTTTCT	584
PI	Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;			
PI	Zhou P, Goodrich R, Drmanac RT;			
XX	WPI: 2001-442253/47.	433	LeuSerAlaGlnSerValMetGluLeuLeuAsnThrAlaProValGlnGluSerProPro	452
DR	P-PSDB; AAM39254.	585	CTGTTCAGCCAGTCTGTGTAATGGAAGAGCTAAACACTGCACCCGTCGAAGAGTCTTAAG	644
XX		453	LeuAlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluVallys	472
PT	Novel nucleic acids and polypeptides, useful for treating disorders such	645	TTGCCATGCTCCTCGGAACCTCACATGGTCTAGAAGTGGCTCATTTGGCTGAAGTTAAG	704
PT	as central nervous system injuries.			
XX	Claim 1; SEQ ID NO 613; 10078pp; English.	473	GluAsnProProPheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGlu	492
CC	The invention relates to human nucleic acids (AAI57798-AAI61369) and the	705	GAGAACCTCTCTTCTATGGGTAATCCGTGGATCGTCAGCCACAGGACTGAATGAA	764
CC	encoded polypeptides (AAM38642-AAM42213) with nootropic,			
CC	immunopressant and cyostatic activity. The polynucleotides are useful	493	ValLeuAlaGlyLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArg	512
CC	in gene therapy. A composition containing a polypeptide or polynucleotide	765	GTCTCGCTCGACTGGAACTGGGAAGATGAGTGTGCGGCTGTACGGATGGAACTTCAGA	824
CC	system, such as peripheral nervous injuries, peripheral neuropathy and			
CC	localised neuropathies and central nervous system diseases, such as	513	GlyThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCys	532
CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic	825	GGCACTCGGTATTTCACTGTGCCCTGAAGAGCGCTGTTTGAAACTGAAGAGCTGC	884
CC	lateral sclerosis, and Shy-drager Syndrome. Other uses include the			
CC	utilisation of the activities such as: Immune system suppression,	533	ArgProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsn	552
CC	Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic	885	AGCCTGACTTAGTGTTCATCATTCAGCCGGTTTCCAATCAGATTGAGCGCTGTAAC	944
CC	and thrombolytic activity, cancer diagnosis and therapy, drug screening,			
CC	assays for receptor activity, arthritis and inflammation, leukaemias and	553	SerLeuAlaPheGlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMet	572
CC	C.N.S disorders. Note: The sequence data for this patent did not form	945	TCCTTAGCATTTGAGGCTACTTAAAGTGAAGTAGTAGAAGAAATACTCCACCAAAATG	1004
CC	part of the printed specification			
XX	Sequence 2523 BP; 757 A; 503 C; 557 G; 706 T; 0 U; 0 Other;	573	GluLysGluGlyLeuGluIleMetIleGlyLysLysLysGlyIleGlnGlyHisTyrAsn	592
SQ	Alignment Scores:	1005	GAAAAAGAGGCTTGGAGATAATGATTGGGAAGAAAGAGCATCCAGGGCTATTACAAAT	1064
	Pred. No.:			
	Score:	1,22e-312	Length:	2523
	Percent Similarity:	3676.00	Matches:	692
	Best Local Similarity:	99.28%	Conservative:	0
	Query Match:	73.02%	Mismatches:	3
	DB:	4	Indels:	2
			Gaps:	2
US-09-671-687A-3 (1-949) x AAI58410 (1-2523)				
QY	255 GlyLysGluSerLeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrp	274		
Db	45 GGAATAAAGCTTTTATATTTTGTGTGGACATGATGATACCTATTGGCAACTGG	104		
QY	275 AspGlyArgPheAspGlyVal----LeuCysSerPheAlaCysValGluSerThrIleLeu	293		
Db	105 GATGGAAGATTTCATGAGGTGCGAGCTTTGTAGTTTTCGGTGTGTGAAAGTACAAATCTA	164		
QY	294 LeuHisIleAsnAspIleIleProGluSerValThrGlnLysArgProProLysLeu	313		
Db	165 TTGCACATCAATGATATATATCCAGAGAGTGTACCGAGAAAGGAGGCTCCCAAACTT	224		
QY	314 AlaPheMetSerArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAla	333		
Db	225 GCCTTTATGTCAAGAGGTGTGGGCAAAAGGTTTCATCCAGTCATATAAACCAGAGCT	284		
QY	334 ThrGlySerThrSerAspProGlyAsnArg--ArgSerGluLeuPheTyrThrLeuAsn	352		
Db	285 ACAGGATCTACCTCAGACCTCGGAATAGAAACAGATCTGAATTTTATACCTTAAT	344		
QY	353 GlySerSerValAspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGlu	372		
Db	345 GGGTCTCTGTTGACTCACAAACACAAATCCAAATCAAAATAACATGATGATGAA	404		
QY	373 ValAlaGluAspProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSer	392		
Db	405 GTTGCAAGAGACCTCGCAAAATCTTTACAGAGATATCTACAGACTTTTGACCGTCTTCA	464		
QY	393 ProProLeuGlnProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeu	412		

QY 773 LeuLeuGluAspThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCys 792
DB 1605 TTACTTTGAAGACACTCCAGACAGTCCCGGATATGTGGAGGGCTTGCAATGTATGAGTGT 1664
QY 793 ArgGluCysTyrAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThr 812
DB 1665 AGAAGATGCTACGACATCCGACATCTCAGCTGGAAATATCAAGCAGATTTTGTAAACC 1724
QY 813 CysAsnThrGlnValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSer 832
DB 1725 TGCAACACTCAAGTCCACCTTCATCCGAGAGGCTGAATCATATAATAACCCAGTGCA 1784
QY 833 LeuProLysAspLeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMet 852
DB 1785 CTTCCCAAGACTTACCCGACTGGGACTGGAGACACGGCTGCATCCCTTGCCCAATATG 1844
QY 853 GluLeuPheAlaValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGly 872
DB 1845 GAGTTATTTGCTGTTCTCTGCATAGAAACAAGCCACTATGTCTTTTGTGAAGTATGG 1904
QY 873 LysAspAspSerAlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsn 892
DB 1905 AAGGACGATCTGCTGGCTCTCTTTGACAGATGGCCGATCGGATGGTGTGAGAT 1964
QY 893 GlyPheAsnIleProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSer 912
DB 1965 GGCTTCAACATTCCTCAAGTCAACCCATGCCAGAGAGTAGGAGTACTTTGAAGATGCT 2024
QY 913 LeuGluAspLeuHisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeu 932
DB 2025 CTGGAAAGACCTGCTATTCCTTTGGACTCCAGGAGATCCAGGCTGTGCAAGACTGCTT 2084
QY 933 CysAspAlaTyrMetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
DB 2085 TGTGATCATATATGTCATGTACAGAGTCCAGAGTCCAAAGTATGTTGACAAA 2135

RESULT 11

ADQ98620
ID ADQ98620 standard; cDNA; 2523 BP.

XX AC ADQ98620;

XX DT 23-SEP-2004 (first entry)

XX DE DNA encoding human GPCR-like protein seqid 290.

XX KW ophthalmological; immunomodulatory; cytostatic; antiatherosclerotic;
XX KW antidiabetic; GPCR-like protein; ophthalmic disorder;
XX KW neurological disorder; immunological disorder; nephritic disorder;
XX KW hormonal dysfunction; cancer; atherosclerosis; diabetes;
XX KW molecular weight marker; food supplement; human; ss.

XX OS Homo sapiens.

XX PN US6569662-B1.

XX PD 27-MAY-2003.

XX PF 19-JUL-2000; 2000US-00620312.

XX PR 21-JAN-2000; 2000US-00488725.

XX PR 25-APR-2000; 2000US-00552317.

XX XX (HYSE-) HYSEQ INC.

XX XX Tang YT, Zhou P, Drmanac RT;

XX XX WPI; 2001-442255/47.

XX New G-protein-coupled receptor-like polypeptides and polynucleotides,
PT useful for treating diseases of ophthalmic, neurological, immunological
PT and nephritic systems and hormonal dysfunction, cancer, atherosclerosis

PT and diabetes.

XX Example 2; SEQ ID NO 290; 92pp; English.

XX The invention describes an isolated polynucleotide (I) comprising a fully
CC defined (SI) of 749, 3188, 2484, 1169, 2936, 1467, 5773, 5714, 4041,
CC 1372, 3996, 3945, 2735, 1788, 585, 1782, 927, 5714 or 2282 nucleotides as
CC given in the specification, its translation or protein coding portion, its
CC extracellular portion or its active domain. The GPCR-like polypeptides
CC and polynucleotides are useful for the treatment of diseases of
CC ophthalmic, neurological, immunological and nephritic systems. They may
CC also be used to treat hormonal dysfunction, cancer, atherosclerosis and
CC diabetes. The antibodies are useful for detecting or quantitating the
CC polypeptide in tissue. The polypeptides can also be used as molecular
CC weight markers and as a food supplement. This sequence represents a human
CC polynucleotide of the invention.

SQ Sequence 2523 BP; 757 A; 503 C; 557 G; 706 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1-22e-312 Length: 2523
Score: 3676.00 Matches: 692
Percent Similarity: 99.28% Conservative: 0
Best Local Similarity: 99.28% Mismatches: 3
Query Match: 73.02% Indels: 2
DB: 5 Gaps: 2

US-09-671-687A-3 (1-949) x ADQ98620 (1-2523)

QY 255 GlyLysGluSerLeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrp 274
DB 45 GGAATAATAAGCTTTTATATTTTGTGTGTGACATGATGATCAACCTATTGGCAACTGG 104

QY 275 AspGlyArgPheAspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeu 293
DB 105 GATGGAAGATTTGATGGAGTGACGCTTTGTAGTTTTCGCTGTTGAAGTACAAATCTGA 164

QY 294 LeuHisIleAsnAspIleIleProGluSerValThrGlnGluArgArgProLysLeu 313
DB 165 TTGCACATCAATGATATATCCAGAGAGTGTCCGAGGAGGAGGAGGCTCCCAACTT 224

QY 314 AlaPheMetSerArgGlyValGlyAspLysGlySerSerSerHisAsnLysProLysAla 333
DB 225 GCCTTTATGTCAAGAGGTGTGGGACAAAGGTTTCATCCAGTCAATATAAACCAGGCT 284

QY 334 ThrGlySerThrSerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsn 352
DB 285 ACAGGATCTACCTCAGACCCCTGGAAATAGAAACAGATCTGAATATTTTATACCTTAAAT 344

QY 353 GlySerSerValAspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGlu 372
DB 345 GGGTCTCTGTGACTCACACACACATCCAAATCAAAAATACATGATGATGATA 404

QY 373 ValAlaGluAspProLysSerLeuThrGluIleSerThrAspPheAspArgSerSer 392
DB 405 GTTGCAGAGAGCCCTGCAAAATCTCTTACAGAGATATCTACAGACTTTTGACCGTCTTCA 464

QY 393 ProProLeuGlnProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeu 412
DB 465 CCACCACTCCAGGCTCTCTCTGGAATCTACACACCCGAGAACACATTCCTCTTTA 524

QY 413 ProPheSerLeuThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSer 432
DB 525 CCATTGATCTCAACCAAGATGCCCATATCCAAATGGAAGTATTGGCCACAGTCCACTTCT 584

QY 433 LeuSerAlaGlnSerValMetGluLeuAsnThrAlaProValGlnGluSerProPro 452
DB 585 CTGTGAGCCAGTCTGTAATGGAAGAGCTAAACACATGCACCCCGTCCAAGAGAGTCCACCC 644

QY 453 LeuAlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLys 472
DB 645 TTGGCCATGCTCTCTGGGAACTCACATGGTCTAGAAGTGGGCTCATTTGGCTGAAGTTAAG 704

QY 473 GluAsnProProPheTyrGlyValIleArgTyrIleGlnProProGlyLeuAenGlu 492
DB 705 GAGAACCCCTCTTCATGGGGTAATCCGTTGGATCGGTGAGCAGCAGGACTGAATGAA 764
QY 493 ValLeuAlaGlyLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArg 512
DB 765 GTGCTCGCTGGACTGGAATCGGAAGATGAGTGTGACGGCTGTACGGATGGAACCTTCAGA 824
QY 513 GlyThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCys 532
DB 825 GGCACCTCGGTATTTACCTCTGCGCTGAAGAGCGCTGTTGTGAAACTGAAGAGCTGC 884
QY 533 ArgProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsn 552
DB 885 AGCCCTGACTCTAGGTTTGCATCATTTGACGCGGTTTCCCAATCAGATTGAGCGCTGTAAC 944
QY 553 SerLeuAlaPheGlyGlyTyrLeuSerGluValValGluAsnThrProProLysMet 572
DB 945 TCTTTAGCATTTGGAGGCTACTTAAGTGAAGTAGTAGAATAATCTCCACCAAAATG 1004
QY 573 GluLysGluGlyLeuGluIleMetIleGlyLysGlyIleGlnGlyHisTyrAsn 592
DB 1005 GAAAAAGAGGCTTGAGATAATGATTGGAGAGAAAGGCATCCAGGGTCATTACAAT 1064
QY 593 SerCysTyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThr 612
DB 1065 TCTTGTACTTACACTCAACCTTATTTCTGCTTATTTGCTTTTATGTTCTGTCTGGACACT 1124
QY 613 ValLeuLeuArgProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeu 632
DB 1125 GTGTACTTAGACCCCAAGAAAGACGATGATGATATTTATTTAGTGAACCCCAAGAGCTA 1184
QY 633 LeuArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIle 652
DB 1185 CTGAGGACAGAAATCTTAATCTCTGAGATATATGATATGTTGTGTCACAAAAT 1244
QY 653 MetLysLeuArgLysIleLeuGluLysValGluAlaSerGlyPheThrSerGluGlu 672
DB 1245 ATGAAACTCAGGAAAAATCTTGAAGAGTGGAGGCTGCATCAGGATTTACCTCTGAAGAA 1304
QY 673 LysAspProGluGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeu 692
DB 1305 AAAGATCCTGAGGAATCTTGAATATCTGTTTCATCATATTTAAGGGTAGAACCTTTG 1364
QY 693 LeuLysIleArgSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMet 712
DB 1365 CTAAATAATAGATCAGCAGTCAAAAGGTACAGATGTTACTTCTATCAAAATTTATG 1424
QY 713 GluLysAsnGluLysValGlyValProThrIleGlnLeuLeuGluThrSerPheIle 732
DB 1425 GAAAAAATGAGAAAGTTGGCGTTCCACAAATTCAGCAGTTGTTAGATGTTCTTTATC 1484
QY 733 AsnSerAsnLeuLysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPhe 752
DB 1485 AACAGTAACTGAAATTTGAGAGGACCATCATCATGTTGATTATTGAGATCGCTCGATTT 1544
QY 753 GlyLysAspPheLysLeuPheLysLysIlePhePheSerLeuLeuLeuAsnIleThrAsp 772
DB 1545 GGAAGAGACTTTAACTATTTAAAAAAATTTTCTCTCTGGAATTAATAATAACAGAT 1604
QY 773 LeuLeuGluAspThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCys 792
DB 1605 TTACTTGAAGACACTCCACAGACTGCGGATATGTGGGGCTTGCATGTATGAGTGT 1664
QY 793 ArgGluCysTyrAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThr 812
DB 1665 AGAGATGCTACGACGATCGGACATCTCAGCTGGAAAAATCAAGCAGATTTTGTAAACC 1724
QY 813 CysAsnThrGlnValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSer 832
DB 1725 TGCAACACTCAAGTCCACCTTCATCGAAGAGGCTGAATCATAAATATATACCCAGTGTCA 1784
QY 833 LeuProLysAspLeuProAspTyrAspTyrArgHisGlyCysIleProCysGlnAsnMet 852

DB 1785 CTTCCCAAGACTTACCCGACTGGGACTGGAGACACGGCTGCATCCCTTGCAGATATG 1844
QY 853 GluLeuPheAlaValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGly 872
DB 1845 GAGTATTTGCTGTTCTCTGCATAGAAACAAGCCACTATGTTGCTTTTGTGAAGTATGG 1904
QY 873 LysAspAspSerAlaTyrLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsn 892
DB 1905 AAGCAGCAATTCGCTGGCTCTTCTTTGACAGCATGGCCGATCGGATGGTGCAGAA 1964
QY 893 GlyPheAsnIleProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSer 912
DB 1965 GGCTTCAACATTCCTCAAGTCAACCCATGCCCAAGTAGGAGAGTACTTGAAGATGCT 2024
QY 913 LeuGluAspLeuHisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeu 932
DB 2025 CTGGAAGACCTGCATTCCTTGGACTCCAGGAGAAATCCAAGGCTGTGCACGAAGACTGCTT 2084
QY 933 CysAspAlaTyrMetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
DB 2085 TGTGATGCATATATGTGCATGTACCAGATGCCAACAATGAGTTGTACAAA 2135
RESULT 12
ADB48380
ID ADB48380 standard; cDNA; 2523 BP.
XX
AC ADB48380;
XX
DT 04-DEC-2003 (first entry)
XX
DE Novel human cDNA SEQ ID NO 290.
XX
KW ss; cancer; neurodegenerative disease; human.
XX
OS Homo sapiens.
XX
FN US2003104529-A1.
XX
PD 05-JUN-2003.
XX
PF 04-JAN-2002; 2002US-00037270.
XX
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
PR 19-JUL-2000; 2000US-00620312.
XX
PA (ZHOU/) ZHOU P.
PA (TANG/) TANG Y T.
PA (LIUC/) LIU C.
PA (ASUN/) ASUNDI V.
PA (DRMA/) DRMANAC R T.
XX
XX Zhou P, Tang YT, Liu C, Asundi V, Drmanac RT;
XX WPI; 2003-678194/64.
XX
XX New polynucleotide, useful for treating diseases e.g., cancer or
PT neurodegenerative diseases.
XX
XX Claim 1; SEQ ID NO 290; 99pp; English.
XX
XX The invention relates to a polynucleotide comprising a sequence given in
CC the specification, or its mature protein-coding portion, or its
CC complement. The polynucleotide is useful for treating diseases e.g.,
CC cancer or neurodegenerative diseases and many others listed in the
CC specification. The present sequence represents a novel human cDNA. Note:
CC The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html?docID=20030104529.
XX
XX Sequence 2523 BP; 757 A; 503 C; 557 G; 706 T; 0 U; 0 Other;

late

Alignment Scores:

Fragment Scores:			
Pred. No.:	1.22e-312	Length:	2523
Score:	3676.00	Matches:	692
Percent Similarity:	99.28%	Conservative:	0
Best Local Similarity:	99.28%	Mismatches:	3
Query Match:	73.02%	Indels:	2
DB:	9	Gaps:	2
US-09-671-687A-3 (1-949) x ADB48380 (1-2523)			
QY	255	GlyLysGluSerLeuGlyTyrPheValGlyValAspMetAspAenProIleGlyAsnTip	274
DB	45	GGAAATAAAGCTTTTATATTTTGGTGTGGACATGATATACCTATTTGGCACTGG	104
QY	275	AspGlyArgPheAspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeu	293
DB	105	GATGGAAGATTGTATGGAGTGCAGCTTTGTAGTTTTCGGTGTGTGAAAGTACAATTCCTA	164
QY	294	LeuHisIleAenAspIleIleProGluSerValThrGlnGluArgArgProProLysLeu	313
DB	165	TTGCACATCAATGATATCATCCACAGAGTGTGACGACGAGAAAGAGCGCTCCCAACTT	224
QY	314	AlaPheMetSerArgGlyValGlyAspLysGlySerSerSerHisAenLysProLysAla	333
DB	225	GCCTTTATGTCAAGAGTGTGGGACAAAGTTTCATCCAGTCATATATAAACCAAGGCT	284
QY	334	ThrGlySerThrSerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsn	352
DB	285	ACAGGATCTACCTCAGACCCCTGGAAATAGAAACAGATCTCAATTTATTTTATCCTTAAAT	344
QY	353	GlySerSerValAspSerGlnProGlnSerLysSerLysAenThrTyrTyrIleAspGlu	372
DB	345	GGGTCCTCTGTGTGACTCACACCAACAATCAAAATCAAAATAATCATGGTACATGGATGAA	404
QY	373	ValAlaGluAspProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSer	392
DB	405	GTTCGACAGACCCCTGCAAAATCTCTTACAGAGATATCTACAGACTTTTGACCGTTCTTCA	464
QY	393	ProProLeuGlnProProProValAsnSerLeuThrThrGluAenArgPheHisSerLeu	412
DB	465	CCACCCTCCAGCCTCCTCCTGTGAATCTCATGCACCAGAGAACAGATTCACATCTTTTA	524
QY	413	ProPheSerLeuThrLysMetProAsnThrAenGlySerIleGlyHisSerProLeuSer	432
DB	525	CCATTCAGTCTCCACAAGATGCCCAATACCNAATGGAAGTATGTCGCACAGTCACATTTCT	584
QY	433	LeuSerAlaGlnSerValMetGluGluLeuAenThrAlaProValGlnGluSerProPro	452
DB	585	CTGTCCGCCAGTCTGTAAATGGAAGAGCTAAACACTGCACCCGTCCCAAGAGAGTCCACCC	644
QY	453	LeuAlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLys	472
DB	645	TTGGCCATCGCTCCTGGGAACTCATGTGTCTAGAGTGGGCTCATTTGGGCTGAAGTTAAG	704
QY	473	GluAenProProPheTyrGlyValIleArgTTrpIleGlyGlnProProGlyLeuAenGlu	492
DB	705	GAGAACCTCTCTTTCTATGGGGTAATCCGTTGGATCGGTACGCCACAGAGCTGAATGAA	764
QY	493	ValLeuAlaGlyLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArg	512
DB	765	GTGCTCGCTGGACTGGAACTGGAAGATGAGTGTGCAGGCTGTACGGATGGAACCTTCAGA	824
QY	513	GlyThrArgTyrPheThrCysAlaLeuIlyAlaLeuPheValLysLeuLysSerCys	532
DB	825	GGCACTCGGTATTTTTCAGCTGTGCCCTGGAAGAAGCGCTGTTTGTGAAACTCAAGAGCTGC	884
QY	533	ArgProAspSerArgPheAlaSerLeuGlnProValSerAenGlnIleGluAtrGysAsn	552
DB	885	AGGCTGACTCTAGGTTTGCATCATTTGCAGCCGGFTTCCAATCAGATTGAGCGCTGTAAAC	944
QY	553	SerLeuAlaPheGlyTyrLeuSerGluValValGluGluAenThrProProLysMet	572
DB	945	TCTTTAGCATTTGGAGGCTACTTAAGTGAAGTAGTAGAAGAAATACTCCACCAAAAATG	1004

QY 933 CysAspAlaTyrMetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
Db 2085 TGTGATGATATATGTCATGACAGAGTCCAAATGAGTTGTACAAA 2135

RESULT 13

AAI60196
ID AAI60196 standard; cDNA; 4286 BP.

XX AAI60196;

XX 22-OCT-2001 (first entry)

XX Human polynucleotide SEQ ID NO 4185.

XX Human; neotropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.

XX Homo sapiens.

XX WO200153312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000WO-US034263.

XX 23-DEC-1999; 99US-00471275.

XX 21-JAN-2000; 2000US-00488725.

XX 25-APR-2000; 2000US-00552317.

XX 20-JUN-2000; 2000US-00598042.

XX 19-JUL-2000; 2000US-00620312.

XX 03-AUG-2000; 2000US-00653450.

XX 14-SEP-2000; 2000US-00662191.

XX 19-OCT-2000; 2000US-00693036.

XX 29-NOV-2000; 2000US-00727344.

XX (HYSE-) HYSEQ INC.

PA Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QH;
PI Zhou P, Goodrich R, Drmanac RT;

XX WPI: 2001-442253/47.

XX P-PSDB; AAM41040.

XX Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.

PS Claim 1; SEQ ID NO 4185; 10078pp; English.

XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC encoded polypeptides (AAM3842-AAM4213) with neotropic,
CC immunosuppressant and cytoskeletal activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression.
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification

SQ Sequence 4286 BP; 1279 A; 779 C; 916 G; 1312 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 7.57e-299 Length: 4286

Score: 3523.00 Matches: 685
Percent Similarity: 97.31% Conservative: 2
Best Local Similarity: 97.03% Mismatches: 8
Query Match: 69.98% Indels: 11
DB: 4 Gaps: 4

US-09-671-687A-3 (1-949) x AAI60196 (1-4286)

QY 255 GlyLysGluSerLeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrp 274

Db 44 GGAATAAATAAGCTTTTATATATTTTGGTGTGGACATGGATAACCCCTATTGGCAACTGG 103

QY 275 AspGlyArgPheAspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeu 293

Db 104 GATGGAGATTTTATGAGGTGACAGTCTTGTAGTTTTCGTGTGTGTGAAGTACAAATCTTA 163

QY 294 LeuHisIleAsnAspIleLeuProGluSerValThrGlnGluArgArgProProLysLeu 313

Db 164 TTGCACATCAATGATATCATCCAGAGAGTGTGACGAGGAAAGAGGCGCTCCCAAACTT 223

QY 314 AlaPheMetSerArgGlyValGlyAspLysGlySerSerSerHisAsnLysProLysAla 333

Db 224 GCCTTTTGTCAAGAGGTGTGGGACAAAGGTTTCATCCAGTCATAATAAACCAAGGCT 283

QY 334 ThrGlySerThrSerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsn 352

Db 284 ACAGGATCTACCTCAGACCCCTGGAAATAGAAACAGATCTGAAATTTATTTATACCTTAAT 343

QY 353 GlySerSerValAspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGlu 372

Db 344 GGCTCTCTGTGACTCACAAACCAATCCAAATCAAAAATAATACATGATGATCATGATGAA 403

QY 373 ValAlaGluAspProAlaLysSerLeuThrGluIleSerThrAspPheAspSerSer 392

Db 404 GTTGACAGAGACCTGCATAATCTCTTACAGAGATATCTACAGACTTTTGACCGTCTTCA 463

QY 393 ProProLeuGlnProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeu 412

Db 464 CCACCACCTCCAGGCTCTCTCTGTGAACTCACTGACCACCGAGAACAGATTCACCTCTTTA 523

QY 413 ProPheSerLeuThrLysMetProAsnThrAsnGlySerIleGlyHisSerProIleSer 432

Db 524 CCATTGAGTCTCACCAGAGTCCCAATACCAATGGAAGATATTGGCCACAGTCCACTTTCT 583

QY 433 LeuSerAlaGlnSerValMetGluLeuAsnThrAlaProValGlnGlnSerProPro 452

Db 584 CTGTGAGCCAGTCTCTGTAATGGAGAGCTAAACACTGCACCCCTCCAGAGAGTCCACCC 643

QY 453 LeuAlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLys 472

Db 644 TTGGCCATGCCTCCTGGAACTCACATGCTTAGAAGTGGGCTCATTTGGCTGAAGTTAAG 703

QY 473 GluAsnProProPheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGlu 492

Db 704 GAGAACCTCTCTTCTATGGGGTAATCCGTGTGATCGGTGAGCCAGCCAGGAGTCAATGAA 763

QY 493 ValLeuAlaGlyLeuGluLeuAspGluCysAlaGlyCys-ThrAspGlyThrPheAr 512

Db 764 GTCTCGTGGATGGAACTGGAAAGATGAGTGTGAGGCTGGTACGGATGGAACCTTCA 823

QY 512 g--GlyThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSer 531

Db 824 GAAGGCACCTCGGTATTTCACCTGTGCCCTCAAGAAGCGGCTGTTTGTGAAACTGAAGAGC 883

QY 532 CysArgProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCys 551

Db 884 TGCAGGCGCTGACTTAGTGTTCATCATTTGACGCGGTTTCCAAATCAGATTCAGGCGGTGT 943

QY 552 AsnSerLeuAla---PheGlyGlyTyrLeuSerGluValValGluGluAsnThrPro--- 569

Db 944 AACTCTTTAGCCATTTGGAGGCGGCTTACTTAAGTGAAGTAGTGGAGGAAATATACCTCCACC 1003

QY 570 ProLysMetGluLysGluGlyLeuGluIleMetIle-GlyLysLysLysGlyIleGlnG1 589

Db 1004 CAAAAATCGGAAAAAGAGGCTTGAGATAATGATTGGGGAAGAAAGGCGATCCAGGG 1063
Qy yHisTyrAenSerCysTyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSerVa 609
Db TCATTACAAATCTTGTTACTTAGACTCAACCTTATCTGCTTATTTGCTTTAGTTCGT 1123
Qy lLeuAspThrValLeuLeuArgProLyseGluLysAenAspValGluTyrTyrSerGluTh 629
Db TCTGACACTGTGTTACTTAGACCCCAAGAAAGAACGATGTAGAATATATATAGTGAAC 1183
Qy rGlnGluLeuLeuArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCysAl 649
Db CCAAGAGCTACTGAGGACAGAAATTTGTAATCCTCTGAGAAATATATGGATATGTGTGTC 1243
Qy aThrIysIleMetIysLeuArgLysIleLeuGluLysValGluAlaAlaSerGlyPheTh 669
Db CACAAAAATTTGAAACTGAGGAAATACTTGAAAGGTGGAGCTGCATCAGGATTTAC 1303
Qy rSerGluGluLysAspProGluGluPheLeuAsnIleLeuPheHisIleLeuArgVa 689
Db CTCTGAAGAAAGATCCTGAGGAATCTTGAATATCTGTTTCATCATATTTTAAGGCT 1363
Qy lGluProLeuLeuLysIleArgSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGl 709
Db AGAACCTTTGCTAAAAATAAGATCAGCAGGTCAAAAGGTACAGATTTGTTACTTCTATCA 1423
Qy nIlePheMetGluLysAenGluLysValGlyValProThrIleGlnGlnLeuLeuGluTr 729
Db AATTTTATGGAAGAAAAATGAGAAAGTTGGCGTCCCAAAATTCAGACGTTGTTAGAAATG 1483
Qy pSerPheIleAenSerAsnLeuLysPheAlaGluAlaProSerCysLeuIleIleGlnMe 749
Db GTCCTTTTATCAACAGTAACTGAAATTTGCAGAGGCACCATCATGCTGATTTATTCAGAT 1543
Qy tProArgPheGlyLysAspPheLysLeuPheLysIleIlePheProSerLeuGluLeuAs 769
Db GCCTCGATTTGAAAGAGCTTTAAACTATTTAAATAATTTTCTCTCTGGAATATAA 1603
Qy nIleThrAspLeuLeuGluAspThrProArgGlnCysArgIleCysGlyGlyLeuAlaMe 789
Db TATAACAGATTTACTTTGAAGACACTCCACAGAGTCCCGGATATGGAGGGCTTCCAAT 1663
Qy tTyrGluCysArgGluCysTyrAspAspProAspIleSerAlaGlyLysIleLysGlnPh 809
Db GTATGAGTGAGAAATGCTACACATCCGACATCTCAGCTGGAAAAATCAAGCAGTT 1723
Qy eCysLysThrCysAsnThrGlnValHisLeuHisProLysArgLeuAsnHisLysTyrAs 829
Db TTGTAAAACTGCAACACTCAAGTCCACCTTCATCCGAGAGGCTGAATCATAAATATAA 1783
Qy nProValSerLeuProLysAspLeuProAspTrpAspTrpArgHisGlyCysIleProCy 849
Db CCCAGTGTCACTTCCCAAGACTTACCAGACTGGGACTGGGACACAGGCTGCATCCCTTG 1843
Qy sGlnAsnMetGluLeuPheAlaValLeuCysIleGluThrSerHisTyrValalalalPheVa 869
Db CCAGAAATATGGAGTATTTGCTGTTCTCTGCATAGAAACAAAGCCACTATGTTGCTTTGT 1903
Qy lLysTyrGlyLysAspAspSerAlaTrpLeuPhePheAspSerMetAlaAspArgAspGl 889
Db GAAGTATGGGAGGACGATCTGCTGGCTCTCTTTTGACAGATGGCCGATCGGGATGG 1963
Qy yGlyGlnAenGlyPheAsnIleProGlnValThrProCysProGluValGlyGluTyrLe 909
Db TGGTCAGAAATGGCTTCAACATTCCTCAAGTCCACCCATGCCAGATAGGAGTACTT 2023
Qy uLysMetSerLeuGluAspLeuHisSerLeuAspSerArgArgIleGlnGlyCysAlaAr 929
Db GAAGATGTCTCTGGAAGAGCTGATCTCTTGGACTCCAGAGAAATCAAGGCTGTGCACG 2083
Qy gArgLeuLeuCysAspAla-TyrMetCys-MetTyr-GlnSerProThrMetSerLeuTyr 948

Db 2084 AGACTGCTTTTGTGATGCCATATATGTCATGTACCCAGAGTCCCAACAATGAGTTGTA 2143
Qy 948 rLys 949
Db 2144 CAAA 2147
RESULT 14
AAS41081
ID AAS41081 standard; cDNA; 2488 BP.
XX AAS41081;
AC AAS41081;
XX 17-DEC-2001 (first entry)
XX cDNA encoding novel human enzyme polypeptide #297.
XX Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
KW ligase; hyperproliferative disorder; immunodeficiency disorder;
KW autoimmune disorder; neurological disorder; metabolic disorder;
KW inflammatory disorder; cardiovascular disorder; reproductive disorder;
KW blood-related disorder; infectious disorder; gene therapy; cytostatic;
KW anti arthritic; nephrotropic; anticoagulant; ss.
XX Homo sapiens.
OS XX
PN WO200155301-A2.
XX 02-AUG-2001.
PD
XX
PF 17-JAN-2001; 2001WO-US001239.
XX
PR 31-JAN-2000; 2000US-017906SP.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225477P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.

QY 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60
DB 366 GTACCAAGGAGTAGTACGACATATATTCAAGATCGTTCTGTGGGGCATTTCAARGATT 425
QY 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGlnProHisAla 80
DB 426 CCTTCTGCAAAAGGCAAGAAATCAATCAGATTGGATTAAATAATCTAGACAACTCATGCA 485
QY 81 ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu 99
DB 486 GTTCTCTTTGTTGATGAAAGGATGTTGTAGAGATAAATGAAAGTTTCACAGAGTTACTT 545
QY 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
DB 546 TTGGCAATTTACCAATTTGTGAGGAGAGGTTTCAGCCTGTTTAAAAACAGARACAGACTAAGT 605
QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
DB 606 AAGGCTCCAAATAGACGTGGCTGTCTGTGAAGTACAGCTGAGATCTGGGGAAGAA 665
QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
DB 666 AATTTCTGGAGTTGTACGCTTCAGAGGACCCCTGTTAGCAGAGGACAGTCTCGGA 725
QY 160 IlePhePheGlyValGluLeuLeuGluGluGlyArgGlyGlnGlyPheThrAspGlyVal 179
DB 726 ATATTTCTTTGGAGTTGAAATGCTGGAAGAGGTCGTGGTCAAGGTTTCACTGACGGGGTG 785
QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
DB 786 TACCAAGGAAACAGCTTTTTCAGTGTGATGAAGATTGTGGCGTGTGTTGTGATTTGGAC 845
QY 199 LysLeuLeuLeuIleGluAspAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
DB 846 AAGCTAGAAGTCTAGAAGATGATGACACTGCAATGGAAAGTATTACGCAGGTCTCTGGG 905
QY 219 AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly 238
DB 906 GACACAATGCAGTGCAGTCTCTCTCTTGGAAATAAATCCAGAGTTTCTTTGAG--- 962
QY 239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
DB 962 ----- 962
QY 259 LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe 278
DB 963 -----GATAACCTTATTGGCAACTGGGATGGAAGATT 995
QY 279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
DB 996 GATGGAGTGCAGCTTTGTAGTTTGGTGTGTTGAAAGTACAATTTCTATTGGCATCAAT 1055
QY 298 AspIleIleProGluSerValThrGlnGluArgProProLysLeuAlaPheMetSer 317
DB 1056 GATATCATCCACAGAGTGTGACGACGAGAAAGGAGGCTCCCAAACTTGCCTTTATGTCA 1115
QY 318 ArgGlyValGlyAspLysGlySerSerSerHisAsnLysProLysAlaThrGlySerThr 337
DB 1116 AGAGGTGTTGGGACCAAGGTTTCATCCAGTCATAATAAACAAGGCTACAGGATCTACC 1175
QY 338 SerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVal 356
DB 1176 TCAGACCTGGAAATGAAACACATCTGAATATTTTATACCTTAAATGGGTCTTCGTT 1235
QY 357 AspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGluValAlaGluAsp 376
DB 1236 GACTCACAACCAACATCCAAATCAAAAATACATGTTGATGATGAAGTTGCAGAGAC 1295
QY 377 ProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerProProLeuGln 396
DB 1296 CCTGCAAAATCTCTTACAGAGATATCTACAGACTTTTGACCGTTCTTCCACCACCACTCCAG 1355

QY 397 ProProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeu 416
DB 1356 CCTCTCTCTGTGAACCTACCTGACCCAGAGAACAGATTCCACTCTTTACCATTCACTC 1415
QY 417 ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGln 436
DB 1416 ACCAAGATGCCCATACCAATGGAGATTGGCCACAGTCCACTTCTCTGTGAGCCAG 1475
QY 437 SerValMetGluGluLeuLeuAsnThrAlaProValGlnGlnSerProProLeuAlaMetPro 456
DB 1476 TCTGTAAATGGAGAGCTAAACACTGCACCGCTCAAAGAGAGTCCACCCTTGGCCATGCCT 1535
QY 457 ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro 476
DB 1536 CCTGGGAACCTACATGCTCTAGAAATGGGCTCATTTGGCTGAAGTTAAGAGAACCTCTCT 1595
QY 477 PheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGly 496
DB 1596 TTTCTATGGGGTAATCCGTTGGATCGGTACGCCACAGGACTGAATGAAGTGTCTCGCTGGA 1655
QY 497 LeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr 516
DB 1656 CTGGAACCTGGAAGATCAGTGTGCAGGCTGTACGATGGAACCTTCAGAGGCACCTCGGTAT 1715
QY 517 PheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSer 536
DB 1716 TTCACCTGTGCCCCGAGAGGCGCTGTTGTGTAACCTGAAGAGCTGCAGGCTGACTCT 1775
QY 537 ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe 556
DB 1776 AGGTTTGGCATCATTTGCAGCGCGTTTCCAAATCAGATTGAGCGCTGTAACCTCTTAGCATTT 1835
QY 557 GlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluGly 576
DB 1836 GGAAGCTACTTAAGTGAAGTAGTAGAAGAAATACTCCACTT----- 1878
QY 577 LeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu 596
DB 1878 ----- 1878
QY 597 AspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArg 616
DB 1878 ----- 1878
QY 617 ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu 636
DB 1878 ----- 1878
QY 637 IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg 656
DB 1878 ----- 1878
QY 657 LysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGlu 676
DB 1878 ----- 1878
QY 677 GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArg 696
DB 1878 ----- 1878
QY 697 SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu 716
DB 1878 ----- 1878
QY 717 LysValGlyValProThrIleGlnGlnLeuLeuGluTrpSerPheIleAsnSerAsnLeu 736
DB 1878 ----- 1878
QY 737 LysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPhe 756
DB 1878 ----- 1878
QY 757 LysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAsp 776

Db 1878 ----- 1878
 QY 777 ThrProArgGlnCysArgIleCysGlyLeuAlaMetTyrGluCysArgGluCysTyr 736
 Db 1878 ----- 1878
 QY 797 AspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGln 816
 Db 1878 ----- 1878
 QY 817 ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAsp 836
 Db 1879 -----CATCCGAGAGGCTGATCATATAATATACCCAGTGTCACTTCCCAAGAC 1929
 QY 837 LeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAla 856
 Db 1930 TTACCCGACTGGGACTGGAGACACGGCTGCATCCCTTGGCAGAAATATGGAGTTATTGCT 1989
 QY 857 ValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspSer 876
 Db 1990 GTTCTCTGCATAGAAACAGCCACTATGTGCTTTTGTGAAGTATGGAGGACGATTCT 2049
 QY 877 AlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlnAsnGlyPheAsnIle 896
 Db 2050 GCTGCTCTCTTTTACAGCATGGCCGATCGGATGGTGGTCAAGTGCCTTCAACATT 2109
 QY 897 ProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLeu 916
 Db 2110 CCTCAAGTCACCCCATGCCAGAGTAGGAGAGTACTTGAAGATGCTCTGGAGACCTG 2169
 QY 917 HisSerLeuAspSerArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyr 936
 Db 2170 CATTCCTTGGACTCCAGGAGATCCCAAGCTGTGCACGAGACTGCTTTGTGATGCATAT 2229
 QY 937 MetCysMetTyrClnSerProThrMetSerLeuTyrIlys 949
 Db 2230 ATGTGCATGTACAGAGTCCCAACATGAGTTGTACAAA 2268

RESULT 15

AAH18478

ID AAH18478 standard; cDNA; 2569 BP.

XX

AC AAH18478;

XX

DT 26-JUN-2001 (first entry)

XX

DE Human cDNA sequence SEQ ID NO:18586.

XX

KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX

OS Homo sapiens.

XX

PN EP1074617-A2.

XX

PD 07-FEB-2001.

XX

PF 28-JUL-2000; 2000EP-00116126.

XX

PR 29-JUL-1999; 99JP-00248036.

XX

PR 27-AUG-1999; 99JP-00300253.

XX

PR 11-JAN-2000; 2000JP-00118776.

XX

PR 02-MAY-2000; 2000JP-00183767.

XX

PR 09-JUN-2000; 2000JP-00241899.

XX

PA (HELI-) HELIX RES INST.

XX

PI Ota T, Isogai T, Nishikawa T, Hayaashi K, Saito K, Yamamoto J;

XX

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX

DR WPI; 2001-318749/34.

XX

PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-

PT length cDNAs defined in the specification, and for the detection and/or
 PT diagnosis of the abnormality of the proteins encoded by the full-length
 XX cDNAs.

PS Claim 8; SEQ ID NO 18586; 2537pp + Sequence Listing; English.

XX
 CC The present invention describes primer sets for synthesizing 5602 full-
 CC length cDNAs defined in the specification, where a primer set comprises:
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesising polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention

XX
 SQ Sequence 2569 BP; 788 A; 470 C; 562 G; 749 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 3.6e-255 Length: 2569
 Score: 3022.50 Matches: 587
 Percent Similarity: 98.66% Conservative: 2
 Best Local Similarity: 98.32% Mismatches: 3
 Query Match: 60.04% Indels: 5
 DB: 4 Gaps: 4

US-09-671-687A-3 (1-949) x AAH18478 (1-2569)

QY 88 ValValGluIleAsnGlnLysPheThrGluLeuLeuAlaIleThrAsnCysGluGlu 107
 Db 1 GTTGTAGAGATAAATGAAAAGTTTCACAGAGTTACTTTTGGCAATTACCAATTGTGAGGAG 60
 QY 108 AtgPheSerLeuPheLysAsnArgAsnArgLeuSerLysGlyLeuGlnIleAspValGly 127
 Db 61 AGGTTGAGCTCTGTTTAAAAACAGAAACAGACTAAGTAAAGCCCTCCAAATAGACGTGGGC 120
 QY 128 CysProValLysValGlnLeuArgSerGlyGluGluLysPheProGlyValValArgPhe 147
 Db 121 TGTCTCTGTGAAGTACAGCTGAGATCTGGGGAAGAAAAATTTCTTGAGTTGTACGCTTC 180
 QY 148 ArgGlyProLeuLeuAlaGluArgThrValSerGlyIlePhePheGlyValGluLeuLeu 167
 Db 181 AGAGGACCCCTGTTAGCAGAGAGCAGACTCTCCGGAATATTTCTTGAGTTGAAATTGCTG 240
 QY 168 GluGluGlyArgGlyGlnGlyPheThrAspGlyValTyrGlnGlyLysGlnLeuPheGln 187
 Db 241 GAAGAAGGTGCTGCTCAAGGTTTCACTGACGGGTGTACCAAGGGAACACAGCTTTTTCAG 300
 QY 188 CysAspGluAspCysGly----PheValAlaLeuAspLysLeuGluLeuIleGluAspAsp 206
 Db 301 TGTGATGAAGATTGTGGCGTGTGTTTGTTCATTGCAAGCTAGCAAGCTCATAGAAGATGAT 360
 QY 207 AspThrAlaLeuGluSerAspTyrAlaGlyProGlyAspThrMetGlnValGluLeuPro 226
 Db 361 GACACTGCATTGGAAAGTGATTACGCGAGTCTCTGGGACACACATGCGAGGTGCAACTTCT 420
 QY 227 ProLeuGluIleAsnSerArgValSerLeuLysGlyGlyGluThrIleGluSerGlyThr 246
 Db 421 CCTTTGAAAATAAATCCAGAGTTCTTTGAAGGTTGGAGAAACAATAGATAATCTGGAAACA 480

```
QY 247 ValIlePheCysAspValLeuProGlyLysGluSerLeuGlyTyrPheValGlyValAsp 266
Db 481 GTTATATTCTGTGATGTTTGCAGAAAGAAAGCTTAGATATTGTTGTGTGGAC 540
QY 267 MetAspAsnProIleGlyAsnTrpAspGlyArgPheAspGlyVal---LeuCysSerPhe 285
Db 541 ATGGATAACCTATTGGCAACTGGGATGGAAGATTGATGGAGTGCAGCTTTCTAGTTT 600
QY 286 AlaCysValGluSerThrIleLeuLeuHisIleAsnAspIleIleProGluSerValThr 305
Db 601 GCCTGTGTTCAAGTAGTCAATCTTATTGTCACATCAATCATATCATCCACAGAGTGTGACG 660
QY 306 GlnGluArgProProLysLeuAlaPheMetSerArgGlyValGlyAspLysGlySer 325
Db 661 CAGGAAGAGGCGCTCCCAAACTTGCCTTTATGTCAGAGGCTTGGGGCAAAAGGTTCA 720
QY 326 SerSerHisAsnLysProLysAlaThrGlySerThrSerAspProGlyAsnArg---Arg 344
Db 721 TCCAGTCATAATAACCAAGGCTACAGATCTACCTCAGACCTCGAATAATAGAACAGA 780
QY 345 SerGluLeuPheTyrThrLeuAsnGlySerSerValAspSerGlnProGlnSerLysSer 364
Db 781 TCTGAATTATTTTATACCTTAATATGGTCTTCTGTGACTCACACCAACCAATCCAATCA 840
QY 365 LysAsnThrTrpTyrIleAspGluValAlaGluAspProAlaLysSerLeuThrGluIle 384
Db 841 AAAAATACATGTTACATGATGAAGTTGCAGAAAGCCCTGCAAAATCTCTTACAGAGATA 900
QY 385 SerThrAspPheAspArgSerSerProLeuGlnProProValAsnSerLeuThr 404
Db 901 TCTACAGACTTTGACCGCTTCTTACCACCACTCCAGCCTCCTCTGTGAACCTCACTGACC 960
QY 405 ThrGluAsnArgPheHisSerLeuProPheSerLeuThrLysMetProAsnThrAsnGly 424
Db 961 ACCGAGACAGATTCACCTCTTACCATTCAGTCTCACCAGAGTGCCCAATACCAATGGA 1020
QY 425 SerIleGlyHisSerProLeuSerLeuSerAlaGlnSerValMetGluGluLeuAsnThr 444
Db 1021 AGTATTGGCCACAGTCCACTTTCTCTGTCCAGCCAGTCTGTATGGAAGAGCTAAACACT 1080
QY 445 AlaProValGlnGluSerProProLeuAlaMetProProGlyAsnSerHisGlyLeuGlu 464
Db 1081 GCACCCGTCACAGAGAGTCCACCTTTGGCCATCCCTCCTCGGSAACCTCACATGGTCTAGAA 1140
QY 465 ValGlySerLeuAlaGluValLysGluAsnProPheTyrGlyValIleArgTrpIle 484
Db 1141 GTGGGCTCATGGCTGAGTTAGGAGAACCTCTCTTCTATGGGGTAATCCGTTGGATC 1200
QY 485 GlyGlnProProGlyLeuAsnGluValLeuAlaGlyLeuGluLeuAspGluCysAla 504
Db 1201 GGTCAGCCACCAGGACTGATGAAGTCTCGCTGGACTGGAACTGGAAGATGAGTGTGCA 1260
QY 505 GlyCysThrAspGlyThrPheArgGlyThrArgTyrPheThrCysAlaLeuLysLysAla 524
Db 1261 GGCTGTACGGATGGAACCTTTCAGAGGCACCTCGGTATTTTCACCTGTGCCCTGGAAGAGCG 1320
QY 525 LeuPheValLysLeuLysSerCysArgProAspSerArgPheAlaSerLeuGlnProVal 544
Db 1321 CTGTTTGTGAACCTGAAGAGCTGCAGGCCCTGACTCTAGGTTTGCATCATTTGACGCGGTT 1380
QY 545 SerAsnGlnIleGluArgCysAsnSerLeuAlaPheGlyGlyTyrLeuSerGluValVal 564
Db 1381 TCCAAATCAGATTGAGCGCTGTAATCTTTAGCATTTTGGAGGCTACTTAAGTGAAGTAGTA 1440
QY 565 GluGluAsnThrProProLysMetGluLysGluGlyLeuGluIleMetIleGlyLysLys 584
Db 1441 GAAGAAAATACTCCACCAAAATGGAAAAAGAGGCTTGGAGATAATGATTGGGAAGAG 1500
QY 585 LysGlyIleGlnGlyHisTyrAsnSerCysTyrLeuAspSerThrLeuPheCysLeuPhe 604
Db 1501 AAAGGCATCCAGGCTCATTAACATCTTGTATTCTTACTTAGACTCAACCTTAATTCGCTTAATT 1560
QY 605 AlaPheSerValLeuAspThrValLeuLeuArgProLysGluLysAsnAspValGlu 624
```

```
Db 1561 GCITTTAGTTCTGTTCTGGACACTGTGTACTTAGACCCCAAGAAAGACGATGTAGAA 1620
QY 625 TyrTyrSerGluThrGlnGluLeuLeuArgThrGluIleValAsnProLeuArgIleTyr 644
Db 1621 TATTATAGTGAAACCCCAAGAGCTACTGAGGACAGAAATTTGTTAATCCTCTGAGAATATAT 1680
QY 645 GlyTyrValCysAlaThrLysIleMetLysLeuArgLysIleLeuGluLysValGluAla 664
Db 1681 GGATATGTGTGTGCCCAAAATTTATGAAACTGAGGAAAAATACTTTGAAAAGGTGGAGGCT 1740
QY 665 AlaSerGlyPheThrSerGluGluLysAspProGluGluPheLeuAsnIle 681
Db 1741 GCATCAGGATTTCCTCTGAAGAAAAAGGT-----GACCATCTTAACTTA 1785
```

Search completed: April 18, 2005, 15:55:04
Job time : 1160 sec

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 18, 2005, 15:26:40 ; Search time 8972 Seconds

(without alignments)
5125.279 Million cell updates/sec

Title: US-09-671-687A-3

Perfect score: 5034

Sequence: 1 MSSGLWSQEKVTSFYWEERI.....RLLCDAYMCYQSPMTSLYK 949

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 2422767955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgn2.1/USPTO.spool.p/US09671687/runat.18042005.115018.17125/app.query.fasta_1.1095
-DB=GenEmbl -OPWT=fastp -SUFFIX=rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09671687 @CNC 1.1.5932 @runat.18042005.115018.17125 -NCFU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sv.*

13: gb.un.*

14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	4983	99.0	3302	6	CQ834224 Sequence
2	4983	99.0	3540	9	BC012342 Homo sapi
3	4983	99.0	5414	9	AB020656 Homo sapi
4	4980	98.9	3302	6	CQ834222 Sequence

5	4971.5	98.8	5371	9	HSA250014
6	4968.5	98.7	3311	6	CQ834226
7	4962	98.6	3480	6	CQ719792 Sequence
8	4942	98.2	4527	6	BD231207
9	4718.5	93.7	4314	10	AK122389
10	4718.5	93.7	4501	10	BC042438
11	4712	93.6	3267	10	BC082001
12	4601	91.4	2845	6	BD160617
13	4601	91.4	2845	6	AX883937
14	4601	91.4	2845	9	AK024348
15	4600.5	91.4	241990	2	AC123449
16	4600.5	91.4	251132	2	AC098162
17	3676	73.0	2523	6	AR338799
18	3433.5	68.2	2341	9	AK000187
19	3022.5	60.0	2569	6	BD160470
20	3022.5	60.0	2569	6	AX883681
21	3022.5	60.0	2569	9	AK024212
22	2930	58.2	1954	9	AK056226
23	2493	49.5	241630	2	AC131881
24	2468	49.0	2116	6	BD135433
25	2468	49.0	2116	6	CQ867781
26	2468	49.0	2116	6	ARS34837
27	2468	49.0	2116	6	AX017216
28	2108.5	41.9	182059	2	AC133651
29	1736	34.5	251957	2	AC126867
30	1662	33.0	2426	10	BC049879
31	1442.5	28.7	212531	2	AC145178
32	1430.5	28.4	129025	2	AC145180
33	1370	27.2	1069	5	BX933038
34	1344	26.7	1335	5	BX934133
35	1262	25.1	837	6	BD149741
36	1262	25.1	837	6	AX869679
37	1215	24.1	758	6	BD147076
38	1215	24.1	758	6	AX867014
39	1132.5	22.5	2068	3	AY071592
40	1111	22.1	151567	2	AC145503
41	1111	22.1	199277	2	AC145445
42	1109.5	22.0	208061	2	AC145254
43	1108.5	22.0	168271	9	AC007728
44	1097.5	21.8	141663	2	AC145018
45	1095.5	21.8	163319	9	HSA303140

ALIGNMENTS

RESULT 1

LOCUS

CQ834224

DEFINITION

Sequence 95 from Patent WO2004058805.

ACCESSION

CQ834224

VERSION

CQ834224.1

KEYWORDS

GI:50833761

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1

Authors

Matsuda A. and Yoneta S.

Title

T cell activating gene

Journal

Patent: WO 2004058805-A 95 15-JUL-2004;

Asahi Kasei Pharma Corporation (JP)

FEATURES

Location/Qualifiers

source

1..3302

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

243..3104

/notes="unnamed protein product"

/codon_start=1

/protein_id="CAH05327.1"

/db_xref="GI:50833762"

/translation="MSSGLWSQEKVTSFYWEERIFYLLQSCVTDKQTKLLKVPKG

SICQYIQDRSVGHSRIPSAKGNKKNIGLKLEQPHAVLFDKVDKVEINEKFTLELLA

ITNCERFSLFKNRNPLSKGLQIDVGCPUKVLQSRSEKFPVVRPRGPILLARTVSG
 IFPGVLLBEGRGQGTGVIYQKQLFCQDEGVFVALDKLELIEDDDTALESVDAG
 PFGVMQVEPLPLEINSRSLKVGETIESGTWFCVLPKGSIGYGVGDMNDPIGNK
 DRGFDQVQCSFACVETILLINDIIPESVTQERRPPKLPFMSRGVGDGKSSHNKP
 KATGSDPNKRNRSLEFVTLNGSSVDSQSKNTWIDEVARDPAKSLTEISDF
 DRSSPLQPPVNSLTTEPNRHSPLSLTKMNTNGSIGHSPLSLSAQSVMEELNAP
 VQESPPLAMPNGSHGLEVSLAEVKENPFVGIWQPPGLNENVLAGELEDSCA
 GCTDGTFRTRVYTCALKKALFVKLSKCRPSRFSALQVSNQIERCNSLAFGGYLS
 VVENETPPKMEEGLEIMTIGKKGIQGHYNSCYLDSLFLFAFSSVLDVLLRPKEK
 NOVEYTSQTELLRLTEINPLRIYGVYCATKMKLRKILEKVEAAGSFTSEKDPPEF
 LNLPHILRLVPELIRAGAKVQDVFVOLFMEKNEKVGVPTIOQLLEWSEINSNL
 KEAEPSCILIIQMPRGKDPKPKKLPSPLELNIITLLEDTPROCRICGLWANYECR
 CDDPDISAGKIQFCKTQNTQVHLHPKRLNKNPVSIPKDLDPDWDRHGHCIPCQNM
 ELFAVLICITSHYVAVKRYKODSALWLFDFSDMRDGGQGNFIPOVTPCPPEVGEYLK
 MSLEDLHSLDSRRIOGCARRLLCDAYMCMYQSPSTMSLYK"

ORIGIN

Alignment Scores:
 Pred. No.: 0 Length: 3302
 Score: 4983.00 Matches: 948
 Percent Similarity: 99.48% Conservative: 0
 Best Local Similarity: 99.48% Mismatches: 1
 Query Match: 98.99% Indels: 4
 DB: 6 Gaps: 4

US-09-671-687A-3 (1-949) x CQ834224 (1-3302)

QY 1 MetSerSerGlyLeuThrPsrGlnGluLysValThrSerProTyrTrpGluGluArgIle 20
 DB 243 ATGAGTTACGGCTTATGGAGCCAAAGAAAGTCACTTCACTTGGGAAAGCGGATT 302
 QY 21 PheTyrLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLys 40
 DB 303 TTTTACTTCTTCAAGATCCAGGCTTACAGACAAACACAAAGCTCCTTAA 362
 QY 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60
 DB 363 GTACCAAGGAGAGTAGACAGTATATTCAAGATCGTTCTGTGGGCACTTCAAGGATT 422
 QY 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGlnProHisAla 80
 DB 423 CCTTTCGAAAAGGCAAGAAATACAGATTGGATTAAATTTAGAGCAACCTCATGCA 482
 QY 81 ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu 99
 DB 483 GTTCTCTTTGTTGATGAAAGGATGTTGTAGAGATAATGAAAGTTTACAGAGTTACTT 542
 QY 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
 DB 543 TTGGCAATTACCAATTGTGAGGAGAGGTTCCAGCTGTGTTAAAAACAGAAACAGACTAAGT 602
 QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
 DB 603 AAAGGCTCCAAATAGACGTGGGCTGTCTGTGAAAGTAGACAGTGTGAGATCTGGGGAAGA 662
 QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
 DB 663 AAATTTCTGGAGTTGTACGCTTCAGAGGACCCCTGTTAGCAGAGAGGACAGTCTCGGA 722
 QY 160 IlePhePheGlyValGluLeuLeuGluGluArgGlyArgGlyGlnGlyPheThrAspGlyVal 179
 DB 723 ATATTCTTTGGAGTTGAAATGCTTGGAAAGAGTTCGTGGTCAAGGTTTCACTGACGGGTG 782
 QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
 DB 783 TACCAAGGAAACACAGCTTTTTCAGTGTGATGAAGATTGTGGCGTGTGTTGTGCAATTGGAC 842
 QY 199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
 DB 843 AAGCTAGAACTCATAGAAGATGATGACACTGCATTGGAAAGTGAATTACGCGAGTCTCTGGG 902
 QY 219 AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly 238

DB 903 GACACAATCGAGGTGCAACTTCTCTCTTGGAAATAACTCCAGAGTTCTTTGAAAGGTT 962
 QY 239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
 DB 963 GGAGNAACAATAGAACTCGAAGTATATCTCTGTGATGTTTTCGACAGAAAGAAAGC 1022
 QY 259 LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe 278
 DB 1023 TTAGGATATTTTGTGTTGGTGTGGACATGATGATAACCTATTGGCAACTGGGATGGAAGATT 1082
 QY 279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
 DB 1083 GATGGAGTGCAGCTTTGTAGTTTGTGCTGTGTTGAAAGTACAATTTATTGGACATCAAT 1142
 QY 298 AspIleIleProGluSerValThrGlnGluArgProProLysLeuAlaPheMetSer 317
 DB 1143 GATATCATCCAGAGAGTGTGACGAGGAAAGAGGGCTCCCAAACTTGGCTTTATGTCA 1202
 QY 318 ArgGlyValGlyAspLysGlySerSerSerHisAsnLysProLysAlaThrGlySerThr 337
 DB 1203 AGAGGTGTTGGGACAAAGGTTTCATCCAGTCAATAATAACCAAGGCTACAGGATCTACC 1262
 QY 338 SerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVal 356
 DB 1263 TCAGACCTCGAAATAGAAACAGATCTGAATTTATACCTTAAATGGGTCTCTGTT 1322
 QY 357 AspSerGlnProGlnSerLysSerLysAsnThrTyrTyrIleAspGluValAlaGluAsp 376
 DB 1323 GACTCAACACCAACATCCAAATCAAAAATACATGTTACATTTGATGAGTTGCAGAGAC 1382
 QY 377 ProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProProLeuGln 396
 DB 1383 CCTGCAAAATCTCTTACAGAGATATCTACAGACTTTTGACCGTTCTTCCACCACCTCCAG 1442
 QY 397 ProProProValAsnSerLeuThrGluAsnArgPheHisSerLeuProPheSerIleu 416
 DB 1443 CCTCTCTCTGTGAACCTCACTGACCCAGAGAAAGAGATTCACACTCTTTTACCATTCAGTCTC 1502
 QY 417 ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGln 436
 DB 1503 ACCAAGATGCCAATACCAGTGAAGTATTTGGCCACAGTCCACTTTCTGTGACGCCAG 1562
 QY 437 SerValMetGluGluLeuLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPro 456
 DB 1563 TCTGTATGGAGAGCTAAACACTGCGCCGCTCCAGAGAGTCCACCCTTGGCCATGCT 1622
 QY 457 ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro 476
 DB 1623 CCTGGGAACCTCACATGCTTAGAAGTGGGCTCATTTGGCTGAAGTTAAGGAGAACCTCTCT 1682
 QY 477 PheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGly 496
 DB 1683 TTTATGGGGTAAATCCGTTGGATCGGTGACCCAGCAGGACTGAATGAAGTGTCTCGCTGGA 1742
 QY 497 LeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr 516
 DB 1743 CTGGAACCTGGAAGATGAGTGTGAGGCTGTACGATGGAACCTTCAGAGGCACTCGGTAT 1802
 QY 517 PheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSer 536
 DB 1803 TTCACCTGTGCTCGGCTGAAGAAGGCGCTGTTGTGAAACTGAAGAGCTGACGCGCTGACTCT 1862
 QY 537 ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe 556
 DB 1863 AGGTTTGGCATCATTTGACGCGGTTTCCAATCAGATTGAGCGCTGTAACTCTTTTAGCATTT 1922
 QY 557 GlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluGly 576
 DB 1923 GGAAGGCTACTTAAGTGAAGTAGTAGAAGAAATACTCCACAAAATGGAAAAAGAGGC 1982
 QY 577 LeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu 596
 DB 1983 TTGGAGATATGATTGGGAAGAAAGAGGATCCAGGCTCATTAACAATTTCTTGTACTTATA 2042

QY 597 AspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuAsg 616
 DB 2043 GACTCAACCTTATCTGCTTATTTGCTTTTAGTTCTGTCGACACTGTGTACTTAGA 2102
 QY 617 ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu 636
 DB 2103 CCCAAGAAAGAACGATGTAGATATATAGTGAACCCAGAGCTACTGAGACAGAA 2162
 QY 637 lleValAsnProLeuArgileTyrGlyTyrValCysAlaThrLysileMetLysLeuAsg 656
 DB 2163 ATTGTTAATCTCTCAGAAATATATGATATGCTGTGCCACAAAATATGAACTGAG 2222
 QY 657 LysIleLeuGluLysValGluAlaSerGlyPheThrSerGluLysAspProGlu 676
 DB 2223 AAAATACTTGAAGAGTGAGGCTGCATCAGGATTTACTCTGAAGAAAAGATCTCTGAG 2282
 QY 677 GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysileAsg 696
 DB 2283 GAATTCCTGAATATCTGTTTCATCATATTTTAAGGTGAGACCTTTGCTAAAATAAGA 2342
 QY 697 SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu 716
 DB 2343 TCAGCAGGTCAAAAGGTACAAGATTGTTACTTCTATCAATTTTATGGAATAATGAG 2402
 QY 717 LysValGlyValProThrIleGlnLeuLeuGluTyrPheIleAsnSerAsnLeu 736
 DB 2403 AAAGTTGGCGTTCCCAACAATTCAGCAGTGTGTAGATGCTCTTTATCAACAGTAACCTG 2462
 QY 737 LysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPhe 756
 DB 2463 AAATTTGAGAGGCGACCATCATGCTGTGATTTATTCAGATCCCTCGATTTGGAAAAGACTTT 2522
 QY 757 LysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAsp 776
 DB 2523 AAATTTTAAATAATTTTCTCTCTGGAATTAATATAACAGATTTACTTTGAGAC 2582
 QY 777 ThrProArgGlnCysArgileCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyr 796
 DB 2583 ACTCCAGACAGTCCGAGATATGTGAGGCGTTCGAATGTATGAGTGTAGAGTAATCTAC 2642
 QY 797 AspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrGln 816
 DB 2643 GACGATCCGGACATCTCAGCTGGAATAATCAAGCAGTTTTGTAAAACCTGCAACACTCAA 2702
 QY 817 ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAsp 836
 DB 2703 GTCCACCTTCATCCGAGAGGCTGAATCATATAATATAACCGATGTCATCTCCAAAGAC 2762
 QY 837 LeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAla 856
 DB 2763 TTACCGAGCTGGAGTGGAGACACGGCTGCATCCCTTGCAGAAATATGAGTTATTTGCT 2822
 QY 857 ValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspSer 876
 DB 2823 GTTCTCTGATAGAAACCAAGCCACTATGTTGCTTTGTGAAGTATGGGAGACGATTTCT 2882
 QY 877 AlaTrpLeuPheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIle 896
 DB 2883 GCCTGGCTCTCTTTGACAGATGCGGATGGGATGGTGGTGCAGAAATGGCTTCAACATT 2942
 QY 897 ProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLeu 916
 DB 2943 CTTCAAGTCAACCCATGCCAGAGTAGAGAGTACTTGAAGATGCTCTGGAGACCTTG 3002
 QY 917 HisSerLeuAspSerArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyr 936
 DB 3003 CATTCCTGGACTCCAGAGAAATCCAAGGCTGTGCACGAGACTGCTTGTGTGATCATAT 3062
 QY 937 MetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
 DB 3063 ATGTGCATGTACAGATGCCAACAATGAGTTGTATACAA 3101

RESULT 2
 BC012342
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 3540)
 Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
 Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
 Altschul,S.F., Zerbberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
 Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
 Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
 Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
 Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
 Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
 Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
 McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
 Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
 Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
 Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,
 Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
 Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
 Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
 Butterfield,Y.S., Krzywinski,M.I., Skalek,U., Smalhus,D.E.,
 Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.,
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002) - late
 22389257
 12477932
 2 (bases 1 to 3540)
 Strausberg,R.
 Direct Submission
 Submitted (15-AUG-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-x@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Rubin Laboratory
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
 BC Cancer Agency, Vancouver, BC, Canada
 info@bcgsc.bc.ca
 Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
 Susanna Chan, Readman Chiu, Chris Fjelli, Erin Garland, Ran Guin,
 Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
 Len Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
 Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeeadi, Jacqueline
 Schein, Duane Smalhus, Michael Smith, Lorraine Spence, Jeff Stott,
 Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
 George Yang, Scott Zuyderduyn, Marco Marra.
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 29 Row: c Column: 5
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 14163257.
 Location/Qualifiers
 1. 3540
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:19923 IMAGE:4552767"
 /tissue_type="Uterus, leiomyosarcoma"

FEATURES
 source

		/clone_lib="NIH_MGC_46"																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																
--	--	-------------------------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

```
Db 1597 TCTGTAATGGAAGAGCTAAACACTGCACCCGCTCCRAAGAGAGTCCACCCCTTGGCCATGCCT 1656
QY 457 ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro 476
Db 1657 CTTGGGAACCTACATGGCTAGAAAGTGGGCTCATTTGGCTGAAGTTAAGGAGAACCCCTCCT 1716
QY 477 PheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGly 496
Db 1717 TTCTATGGGGTAATCCGTTGGATCGGTACGCCACACAGGACTGAATGAAGTCTCGCTGGA 1776
QY 497 LeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr 516
Db 1777 CTGGAACCTGGAAGATGAGTGTGCAGGCTGTACGATGGAACCTTCAGAGGCACTCGGTAT 1836
QY 517 PheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSer 536
Db 1837 TTCACCTGTGCGCTCAAGAAGGCGCTGTTTGTGAACCTGAAGAGCTGCAGGCGCTGACTCT 1896
QY 537 ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe 556
Db 1897 AGGTTTGCATCATTCAGCCGGTTTCCAATCAGATTGAGCGCTGTAACTCTTTAGCATTT 1956
QY 557 GlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluGly 576
Db 1957 GGAGGCTACTTAAGTGAAGTAGTAGAAGAAATACTCCACCAAAAATGGAAAAAGAGGC 2016
QY 577 LeuGluIleMetIleGlyLysLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu 596
Db 2017 TTGGAGATAATGATTGGGAAGAAGAACGATCCAGGGTCATTACAAATCTCTGTACTTA 2076
QY 597 AspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArg 616
Db 2077 GACTCAACCTTATCTGCTTATTTGCTTATTTAGTTCTGTTCGACACTGTGTACTTAGA 2136
QY 617 ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu 636
Db 2137 CCCAAGAAAAAGACGATAGAAATATTATAGTGAACCCAGAGCTACTGAGGACAGAA 2196
QY 637 IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg 656
Db 2197 ATTGTTAATCCTCTGAGATATATGATATGATGTGTGCCACAAAAATTTAAGAACTGAGG 2256
QY 657 LysIleLeuGluLysValGluAlaSerGlyPheThrSerGluLysAspProGlu 676
Db 2257 AAAATACCTTGAAGAGGTGGAGGCTGCATCAGGATTTACCTCTGAAGAAAAAGATCCTGAG 2316
QY 677 GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLysIleArg 696
Db 2317 GAATTTCTTGAATATCTGTTTCATCATATTTTAAAGGGTAGAACCTTTTGCTAAAAAATAAGA 2376
QY 697 SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu 716
Db 2377 TCAGCAGGTCAAAAGGTACAGATTGTTACTTCTATCAAAATTTTATGAAAAAATAGAG 2436
QY 717 LysValGlyValProThrIleGlnGlnLeuLeuGluTrpSerPheIleAsnSerAsnLeu 736
Db 2437 AAAGTTGGCGTCCCAACATTCAGCAGTTGTTAGAAATGGTCTTTTATCAACAGTAACCTG 2496
QY 737 LysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPhe 756
Db 2497 AAAATTTGACAGGCAACCATCTCTGATTAATTCAGATGCTCGATTTGGAAAAAGACTTT 2556
QY 757 LysLeuPheLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAsp 776
Db 2557 AAACATTTTAAAAAAATTTTCTCTCTGGAATTAATAATACACATTTACTTTGAAGAC 2616
QY 777 ThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyr 796
Db 2617 ACTCCACACAGTCCCGATATGTGAGGCGCTTGCAATGTATGAGTGTAGAGATGCTAC 2676
QY 797 AspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGln 816
Db 2677 GACGATCCGACATCTCAGCTGGAAAAATCAAGCAGTTTTTTGTAACCTTGCACACTCAA 2736
```

```
QY 817 ValHisIeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAsp 836
Db 2737 GTCACCTTTATCCGAGAGGCTGAATCATAAATATATAACCCAGTGTCACTTCCCAAGAC 2796
QY 837 LeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAla 856
Db 2797 TTACCCGACTGGGACTGGAGACACGGCTGCATCCCTTGCAGAAATATGGAGTTATTGCT 2856
QY 857 ValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSer 876
Db 2857 GTTCTCTGCATAGAAAACAGCCACTATGTTGCTTTTGTGAAGTATGGAGAGCAGATTCT 2916
QY 877 AlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIle 896
Db 2917 GCCTGCTCTTCTTTGACGATGGCCGATCGGATGGTGGTCAAGATGGCTTCAACATT 2976
QY 897 ProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLeu 916
Db 2977 CCTCAAGTCACCCATCCAGAGTAGGAGAGTACTTTGAAGATGCTCTCTGGAAGACCTG 3036
QY 917 HisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyr 936
Db 3037 CATTCCTTGGACTCCAGGAGATCCAAAGGCTGTGCACGAAGACTGCTTTGTGATGCATAT 3096
QY 937 MetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
Db 3097 ATGTGATGTATCCAGAGTCCAAACAATGAGTTTGTACAAA 3135
```

RESULT 3

AB020656 5414 bp mRNA linear PRI 17-MAY-2001
Homo sapiens mRNA for KIAA0849 protein, partial cds.

LOCUS

AB020656

DEFINITION

AB020656

ACCESSION

AB020656.2

VERSION

GI:14133220

KEYWORDS

Homo sapiens (human)

SOURCE

Homo sapiens

ORGANISM

Homo sapiens

REFERENCE

1 (bases)

AUTHORS

Nagase, T., Ishikawa, K., Suyama, M., Kikuno, R., Hirose, M., Miyajima, N., Tanaka, A., Kotani, H., Nomura, N. and Ohara, O.

TITLE

Prediction of the coding sequences of unidentified human genes. XII. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro

JOURNAL

DNA Res. 5 (6), 355-364 (1998)

MEDLINE

99156230

PUBMED

10048485

REFERENCE

2 (bases 1 to 5414)

AUTHORS

Ohara, O., Suyama, M., Kikuno, R., Nagase, T. and Ishikawa, K.

TITLE

Direct Submission

JOURNAL

Submitted (02-DEC-1998) Osamu Ohara, Kazusa DNA Research Institute, Laboratory of DNA Technology, Yata 1532-3, Kisarazu, Chiba 292-0812, Japan (E-mail: cdnainfo@kazusa.or.jp, Tel: +81-438-52-3913, Fax: +81-438-52-3914)

COMMENT

On May 17, 2001 this sequence version replaced gi:4240186.

FEATURES

Location/Qualifiers

1..5414

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="fh04363"

/sex="male"

/tissue_type="brain"

/clone_lib="pBluescriptII SK plus"

/dev_stage="adult"

/notes="vector:pBluescriptII SK plus. This sequence is replaced that of hk05904 cDNA as a representative cDNA sequence for KIAA0849"

1..5414

/genes="KIAA0849"

<426..3308

gene

CDS

NOT LATE

```

/gene="K1AA0849"
/codon_start=1
/product="K1AA0849 protein"
/protein_id="BA074872.2"
/db_xref="GI:14133221"
/translation="PFVFNITSSGLNSQEKVTSYWEERIFYLILLOBCSVTDKQTK
LKVPKSGIQYIQRSGVHSRIPSAKGNKQIGLILQPHAVLFLVDEKQVVEINEK
FTLELLATNCEERESLFKNRNLKGLQIDVGCYKQVLRSGEEKFGVVRFGPLL
AERTVSGIFFGVLEELRGGQFTDGVYQKQFQCDQEDCGFVALDKLELIEDDTA
LESYVAGPDTMOVBLPPLNRSVSLKGETIESGTVIFCDVLPKESLSGVFVGDM
DIPGNMGRDPGVOLCPACVSESFTLLHINDIIPESVTOERPPKLAAPMSRGVGDGK
SSHNKPKATSGTSIPGNRNSLEFYLTLNGSSVDSQPKSKNWTYIDEVADPAKSL
TEISDFRSPSPLOPPPNLTTNRFHSLPFLTKMPNTNGSIHGSPLSLASQSV
BELNTAQVQESPLPMPGNHSLGVSLAEKPNPFYVIRWIGQPGLENEVLNKL
ELEDECACTDCTFGTRYFTCALKKALFVKLSCRPSRPFASLOPVNOIERCNLA
FGYLSVEVEENTPYSKMEKGLIEMIGKKGIQGHVNSCYLDSITLFCIFAPSSVLDTV
LLRPKEKNDVEYSETOELLREINPLRIYGVCAATKIMKRLKILEKVEAASGTISE
EKDPEFLNPFHHLIRLEPLUKISAGQKQVDCYFIQIFMEKNKGVPTTIOQLLEW
SFINSNLFAPAEAPSLIITQMPRFKDFLFFKXIFPSLELNTDLLLEDTPRQCRICGGL
AMYECECYDDPDIDGAKIKQPKCTCNTQVHLKRLNKHYPVSLPKDLDPDWRHG
CTPCQNMELFVLCITESHYVAFVKYKODSALWLPFDSMADRDGGQNGNIPQVTPCP
EVGEYLKMSLEDLHSLDSRRIQGCARRULLCDAYMCYQSPMTSLYK"
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 5414
Score: 4983.00 Matches: 948
Percent Similarity: 99.48% Conservative: 0
Best Local Similarity: 99.48% Mismatches: 1
Query Match: 98.99% Indels: 4
DB: 9 Gaps: 4
US-09-671-687A-3 (1-949) x AB020656 (1-5414)
QY 1 MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTrpTrpGluGluArgIle 20
DB 447 ATGAGTTACAGGCTTATGAGCCCAAGAAAAAGTCACTTCACTCCCTTCTGGAAGAGCGGATT 506
QY 21 PheTyLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLeuLys 40
DB 507 TTTTACTTGCTTCTTCAAGAATGCAGCGTTACAGACAAAACAAACAAAGCTCTCTTAAA 566
QY 41 ValProLysGlySerIleGlyGlnTyrlleGlnAspArgSerValGlyHisSerArgIle 60
DB 567 GTACCAAGGGAAGTATAGACAGTATATTCAAGATCGTTCCTGTGGGCATTCACGATT 626
QY 61 ProSerAlaLysGlyLysLysAenGlnIleGlyLeuLysIleLeuGlnProHisAla 80
DB 627 CCTTCTGCAAAAGGCAAGAAATATCAGATTGGATTAAAAATTTAGAGCAACCTCATGCA 686
QY 81 ValLeuPheValAspGlu---AspValValGluLeuAsnGlnLysPheThrGluLeuLeu 99
DB 687 GTTCTCTTTGTGATGAAAGGATGTTGTAGACATAAATGAAAGTTACACAGATTACTT 746
QY 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAenArgLeuSer 119
DB 747 TTGGCAATTACCAATTGTGGAGAGAGGTTGAGCTGCTGTTTAAAAACAGAAACAGACTAAGT 806
QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
DB 807 AAAGGCTTCAATATAGACGTGGGCTGTCTGTGAAAGTACAGCTGAGATCTGGGGAAGAA 866
QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
DB 867 AAATTTCTTGAGTTGTGACCTTTACAGAGGACCCCTGTTAGCAGAGAGACAGCTCCGGA 926
QY 160 IlePhePheGlyValGluLeuLeuGluGluArgGlyGlnGlyPheThrAspGlyVal 179
DB 927 ATATTCTTTGGAGTTGAAATGCTGGAAAGAGGTCGTGGTCAAGTTTCACTGACGGGGTG 986
QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
DB 987 TACCAAGGGAAACAGACTTTTTCAGTGTGATGAAGATTGTGGCGGTGTTGTTGTCATTGGAC 1046

```

```

QY 199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
DB 1047 AAGCTAGAACTCATAGAAAGATGATGACATGCTGCTGAAAGTGATTACGCGAGTCTCTGGG 1106
QY 219 AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly 238
DB 1107 GACACAATCGAGGTGAACTTCTCTCTTGGAAATAAATCCAGAGTTCTTTTGAAGGTT 1166
QY 239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
DB 1167 GAGAAACAATAGATCTGCAACAGTTATATTCTGTGATGTTTGGCAGAGAAAAGAAGC 1226
QY 259 LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe 278
DB 1227 TTAGGATAATTTTGTGTTGGTGGACATGGATAACCTTATTGGCAACTGGGATGGAAGATT 1286
QY 279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
DB 1287 GATGGAGTCAGCTTTTGTAGTTTGTGCTGTGTTGAAAGTACAATTTCTATTGCAATCAAT 1346
QY 298 AspIleIleProGluSerValThrGlnGluArgProProLysLeuAlaPheMetSer 317
DB 1347 GATATCATCCAGAGAGTGTGCGCAGGAAAGAGGCGCTCCCAAACTTGCCTTTATGTCA 1406
QY 318 ArgGlyValGlyAspLysGlySerSerSerHisAsnLysProLysAlaThrGlySerThr 337
DB 1407 AGAGGTGTTGGGACAAAGGTTTCATCCAGTCATAATAAACAAGGCTACAGGATCTACC 1466
QY 338 SerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVal 356
DB 1467 TCAGACCTCGGAAATAGAAAACAGATCTGAATTTATTTATACCTTAAATGGGTCTCTGTT 1526
QY 357 AspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGluValAlaGluAsp 376
DB 1527 GACTCAACACCAATCCAAATCAAAAATACATGTTGATGATGATGATGATGATGATGATGAT 1586
QY 377 ProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProProLeuGln 396
DB 1587 CTTGCAAAATCTCTTACAGAGATATCTACAGACTTTTGACCGTCTCTTCAACCACTCCAG 1646
QY 397 ProProProValAsnSerLeuThrGluAsnArgPheHisSerLeuProPheSerLeu 416
DB 1647 CCTCTCTCTGTGAACCTCACTGACACCGAGAACAGATTCACACTCTTTACCACTTCCAGTCTC 1706
QY 417 ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGln 436
DB 1707 ACCAAGATGCCCAATACCAATGGAAGTATTGGCCACAGTCCACTTCTCTGTGAGCCAG 1766
QY 437 SerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPro 456
DB 1767 TCTGTAATGGAAGAGCTAAACACTGACCCGTCAGAGAGAGTCCACCTTGGCCATGCCT 1826
QY 457 ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro 476
DB 1827 CCTGGGAACCTCACATGCTCTAGAAGTGGGCTCATTTGGCTGAAAGTAAAGAGAACCTCTCT 1886
QY 477 PheTyGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGly 496
DB 1887 TTTCTATGGGGTAATTCCTGATCGGTTCAGCCACAGGACTGAAATGAAGTCTCGCTGGA 1946
QY 497 LeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr 516
DB 1947 CTGGAATCGGAAGATGATGTCAGGCTGTACGGATGTAACCTTCAGAGGACCTCGGTAT 2006
QY 517 PheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSer 536
DB 2007 TTTCACTGTCCCTGGAAGAGGCGCTGTTGTGAAACTGAAAGAGCTGCAGGCTGACTCT 2066
QY 537 ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe 556
DB 2067 AGGTTTTCATCATTTGACGCGGTTTCCAAATCAGATTGAGCGCTGTAACCTCTTTAGCATTT 2126

```


557 GlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluGly 576
2127 GGAGGCTACTTAAGTAGAGTAGAGAAATACTCCACCAAAATGGAAAAGAGGC 2186
577 LeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu 596
2187 TTGGAGATAATGATTGGGAAGAAGAGGATCCAGGGTCATTACAAATCTTGTACTTA 2246
597 AspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArg 616
2247 GACTCAACCTTAATCTGCTTATTGCTTTAGTTCCTGTCGACACTGTGTACTTAGA 2306
617 ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu 636
2307 CCCAAGAAGAAAGACGATGAGATATTATAGTGAACCCCAAGAGCTACTGAGACAGAA 2366
637 IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg 656
2367 ATTGTTAATCTCTGAGATATATGATATGCTGTGTCACAAAAATTTATGAACCTGAGG 2426
657 LysIleLeuGluLysValGluAlaSerGlyPheThrSerGluLysAspProGlu 676
2427 AAATATCTTGAAGAGTGAGGCTGCATCAGGATTTACCTCTGAGAAAGATCCTGAG 2486
677 GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArg 696
2487 GAAATCTTGAATATCTGTTTATCATATATTTAAGGGTAGAACCCTTGTCTAAAAATAAGA 2546
697 SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu 716
2547 TCAGCAGGTCAAAAGGTACAGATTGTTACTTCTATCAATTTTATGAAAAAATGAG 2606
717 LysValGlyValProThrIleGlnLeuLeuGluTyrSerPheIleAsnSerAsnLeu 736
2607 AAAGTTGGCGTTCCCAAAATTCAGCAGTTGTTAGAAATGCTCTTTATCAACAGTAACCTG 2666
737 LysPheAlaGluAlaProSerCysIleuIleIleGlnMetProArgPheGlyLysAspPhe 756
2667 AAATTTGAGAGGACCATCATGCTCTGATATTATTCAGATCCCTCGATTTGGAAAAAGCTTT 2726
757 LysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAsp 776
2727 AAATATTAAAAAATTTTCTCTCTGGAATTAATAATACAGATTACTTTGAAGAC 2786
777 ThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyr 796
2787 ACTCCAGACAGTCGGGATATGTCGAGGCTTGCAATGTCATGAGTGTAGAGATGCTAC 2846
797 AspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGln 816
2847 GACGATCCGGACATCTCAGCTGGAAAAATCAAGCAGTTTTTGTAAAAACCTGCAACACTCAA 2906
817 ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAsp 836
2907 GTCCACCTTCATCCGAAGAGGCTGAATCAATAATAATACCCAGTGTCATCTCCCAAGAC 2966
837 LeuProAspTyrAspTyrArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAla 856
2967 TTACCCGACTGGGACTGGAGACACGGCTGCATCCCTTGCAGAAATGAGTATTGTCT 3026
857 ValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspSer 876
3027 GTTCTCTCATAGAAACCAAGCCACTATGTTGCTTTTGTGAAGTATGGGAAGACGATCT 3086
877 AlaTyrPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIle 896
3087 GCCTGGCTTCTTTTGAAGCATGCGGATCGGATGGTGGTCAGAAATGGCTTCAACATT 3146
897 ProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLeu 916
3147 CCTCAAGTACCCCATGCCAGAGTAGAGAGTACTTGAAGATGCTCTGGAAGACCTG 3206
917 HisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAlaTyr 936

3207 CATTCCTTGACTCCAGAGATCCAGGCTGTGCACAGAGCTCTTTGTGATGATAT 3266
937 MetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
3267 ATGTGCAATGTACCAGAGTCCAAACAATGAGTTGTACAAA 3305
RESULT 4
CQ834222 3302 bp DNA linear PAT 29-JUL-2004
LOCUS Sequence 93 from Patent WO2004058805.
DEFINITION CQ834222
ACCESSION CQ834222.1 GI:50833759
VERSION CQ834222.1
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Matsuda, A. and Yoneta, S.
TITLE T cell activating gene
JOURNAL Patent: WO 2004058805-A 93 15-JUL-2004; -late
Asahi Kasei Pharma Corporation (JP)
FEATURES
source 1..3302
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
243..3104
/note="unnamed protein product"
/codon_start=1
/protein_id="CAH05326.1"
/db_xref="GI:50833760"
/translations="MSGSLWSQKVTSPYWEERIFYLLQBCSVTDKOTOKLLKVPKG
SIGQYIQDSVGHSHRIPSAKKNOIGLKILEOPHVLVDEKDVVEINEXFETLLA
IINCEERFSLFKNRLSKGLQIDVGCPVKQLRSGEKFPVVRFRGPLLAERTVSG
IFPVELLEGQGGFTGVYQKQLFQDDEDCGVFVALDKLELIEDDDTALSDYAG
PDGMQVLPPLLEINRSVLKGETIEGTVI FCDVLPKESLGYFVGVDMSHNGW
DGRFDVQLCSFACVESTILLHINDIIPESVTOERRPKLAFMRSGVDGKSGSHNKP
KATGSTDPGNRNSLFTYLLNGSSVDSQPSKSKNTWYIDEAEDPAKSTETSTDF
DRSSPLOPPVNSLTENEFHSLPSTLTKMNTNGSIGHSPLISAQSVMEELNTP
VQESPLPAPPGNSHGLEVSLAEVKENPEFYGVIRWIGQPPGNEVLAGELEDECA
GCTDTPRTGYFTCAKKAFLVKLSRSPDSFASIQVSNQERNCNSLAFGGYLS
VVENTPPKMEKEGLEIMIGKKGIQGHNSCYLDSLFCFLFAFSVLDTVLLRPEK
NDVEYSEQLLRTINPLRYGVCAIKMLKLEKVEAASGFTSEKQDPEEF
LNLPHILRVEPLLIKRSAGQVQDCYFQIEMEKNEKVGVPITIQQLLEWFTNSNL
KFAEAPSCILLIQMPREGKDKPLFKIPPSLELNTDLLEDTPRQCRICGGLAMECRE
CTDDPDISAGIKQPCNTQVHLHPKRLNHNKTPVSLPKDLDFDMDWRHGCICQNM
ELFAPVICITSHYVAFVKYKDSAWLFFDSMDRGGGNGFNIPQVTPCPPEVGEYLK
MSLEDLHLSRRIOGRCARRLLCDAYMCMTQSPMTSLYK"

Alignment Scores:
Pred. No.: 0 Length: 3302
Score: 4980.00 Matches: 947
Percent Similarity: 99.48% Conservative: 1
Best Local Similarity: 99.37% Mismatches: 1
Query Match: 98.93% Indels: 4
DB: 6 Gaps: 4

US-09-671-687A-3 (1-949) x CQ834222 (1-3302)

QY 1 MetSerSerGlyLeuTyrPsrGlnGluLysValThrSerProTyrTrrGluGluArgIle 20
Db 243 ATGAGTTTCAGCTTATGGAGCCCAAGAAAAGTCACTTACCCTACTGGGAAGACGGATT 302
QY 21 PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLys 40
Db 303 TTTTACTTGTCTTCTCAAGATGACAGGTTACAGACAAACAAACACAAAGCTCTCTTAA 362
QY 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60
Db 363 GTACCGAAGGAGTAGTAGACAGTATATTCAAGATCGTTCTGTGTGGGCGCATTCAGGATT 422

QY	61	ProSerAlaIysGlyLysIysAsnGlnIleGlyLeuLysIleLeuGluInProHisAla	80	QY	417	ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGln	436
DB	423	CCTTCTGCAAAAGGCAAGAAATCATGATTGGATTAAAAATTTAGAGCAACCTCATGCA	482	DB	1503	ACCAAGATCCCAATACCAATGGAGTATTGGCCACAGTCCACTTCTCTGTGAGCCAG	1562
QY	81	ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu	99	QY	437	SerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPro	456
DB	483	GTTCCTTTTGTGATGAAAGGATGTTGAGAGATAAATGAAAGTTTCACAGAGTTACTTT	542	DB	1563	TCGTGTAATGGAGAGCTAAACACTGACCCTGCTCAAGAGAGTCCACCTTGGCCATGCT	1622
QY	100	LeuAlaIleThrAsnCysGluGluAArgPheSerLeuPheLysAsnArgAsnArgLeuSer	119	QY	457	ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro	476
DB	543	TTGGCAATTTACCAATTTGAGGAGAGGTTTCAGCTGTTTAAAAACAGAAACAGACTAAGT	602	DB	1623	CCTGGGAACCTCAGATGGTCTAGAGTGGGCTCATTTGGCTGAAGTTAAGAGAACCTCTCT	1682
QY	120	LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu	139	QY	477	PheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGly	496
DB	603	AAAGGCTCCAAATAGACGTGGGCTGCTCTGTGAAAGTACAGCTGAGATCTGGGGAGAA	662	DB	1683	TTCTATGGGGTAAATCCGTTGGATCGGTGAGCCACAGGACTGAATGAAGTGTCTCGCTGA	1742
QY	140	LysPheProGlyValValIleArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly	159	QY	497	LeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr	516
DB	663	AAATTTCTGGAGTTGTACGCTTCAGAGGACCCCTGTTAGCAGAGGACAGTCTCCGGA	722	DB	1743	CTGGAACTGGAGATGAGTGTGAGGCTGTACGGATGGAAACCTTCAGAGGCACCTCGGTAT	1802
QY	160	IlePhePheGlyValGluLeuLeuGluGlyArgGlyGlnGlyPheThrAspGlyVal	179	QY	517	PheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSer	536
DB	723	ATATCTTTGGAGTTGAATGCTGGAGAGAGGTCGTGGTCAAGTTCCTGACGGGGTG	782	DB	1803	TTCACTGTGCTGAGAGAGGCGCTGTTGTGAAACTGAAGAGCTGCAGGCCCTGACTCT	1862
QY	180	TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp	198	QY	537	ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe	556
DB	783	TACCAAGGAAACAGCTTTTTCAGTGTGATGAAGATTGTGGCGTGTGTTGTCATTGGAC	842	DB	1863	AGGTTTGCATCATTTGCAGCGGTTTCCAATCAGATTGAGCGCTGTAACCTCTTTAGCATTT	1922
QY	199	LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly	218	QY	557	GlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluGly	576
DB	843	AAAGCTAGAACTCATAGAAGATGATCACCTGCATTGGAAAGTGATTACGCAGGTCCTGGG	902	DB	1923	GGAGGCTACTTAAGTAGAGTAGAAGAAATACTCCACCAAAATGGAAGAAAGAGGC	1982
QY	219	AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly	238	QY	577	LeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu	596
DB	903	GACCAATGCAGGTCGAATCTCTCTTTGGAAATAAATCCAGAGTTCTTTTGAAGGTT	962	DB	1983	TTGGAGATTAATGATTTGGAGAGAAAGAGGATCAGGCTCATTTACAATTTCTGTACTTA	2042
QY	239	GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer	258	QY	597	AspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArg	616
DB	963	GGAGAAACAATAGAACTGGACAGATTATATCTGTGATGTTTTCAGGAGAAAGAAAGC	1022	DB	2043	GACTCAACCTTATCTGCTTATTTGCTTTTGTGTTTCTGTTCTGGACACTGTGTTACTTAGA	2102
QY	259	LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe	278	QY	617	ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu	636
DB	1023	TTAGGATATTTTGTGTGGACATGATTAACCTATTGGCACTGGGATGGAAGATT	1082	DB	2103	CCCAAGAAAGAAAGACGATGTAGAATATATATAGTGAACCCCAAGAGCTACTGAGACAGAA	2162
QY	279	AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn	297	QY	637	IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg	656
DB	1083	GATGGAGTGCAGCTTTGTAGTTTGGCTGTGTTGAAAGTACAATCTTATGGCAATCAAT	1142	DB	2163	ATTGTTAACTCTCGAATATATGATGATGTGTGTGCCACAAAATAATATGAAGACTGAAG	2222
QY	298	AspIleIleProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMetSer	317	QY	657	LysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGlu	676
DB	1143	GATATCATCCAGAGAGTGTGACGAGGAAAGAGGCGCTCCCAACTTGTGCTTTATGTCA	1202	DB	2223	AAAATACTTGAAAGGTGGAGGCTGCATCAGGATTTACCTCTGAAGAAAAAGATCCTGAG	2282
QY	318	ArgGlyValGlyAspLysGlySerSerSerHisAsnLysProLysAlaThrGlySerThr	337	QY	677	GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArg	696
DB	1203	AGAGGTGTTGGGACAAAGGTTTCATCCAGTCATAATAAACAAGGCTACAGGATCTACC	1262	DB	2283	GAATCTTCGAATATCTGTTTCATCATATTTAAAGGGTAGAACCTTTGCTAAAAAATAAGA	2342
QY	338	SerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVal	356	QY	697	SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu	716
DB	1263	TCAGACCCCTGGAAATGAACAGATCTGAATATTTTATACCTTAAATGGGCTCTCTGTT	1322	DB	2343	TCAGCAGGTCAAAAGGTACAAAGATTGTTACTTCTATCAAAATTTTATGGAAAAAATGAG	2402
QY	357	AspSerGlnProGlnSerLysAsnThrTyrTyrIleAspGluValAlaGluAsp	376	QY	717	LysValGlyValProThrIleGlnLeuLeuGluTrpSerPheIleAsnSerAsnLeu	736
DB	1323	GACTCACAACCAATCCAATCAAAAATATCATGTTGATGATGATGATGATGATGATGATGAT	1382	DB	2403	AAAGTTGGCGTCCCACAATTCAGCAGTGTCTAGATGGTCTTTTATCAACAGTAACTG	2462
QY	377	ProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProProLeuGln	396	QY	737	LysPheAlaGluAlaProSerCysLeuIleGlnMetProArgPheGlyLysAspPhe	756
DB	1383	CCTGCAAAATCTTACAGAGATATCTACAGACTTTTGACCGTCTTCACCCACCTCCAG	1442	DB	2463	AAATTTGCAGAGGCACCATCATGCTGATTAATTTCAGATGCTCGATTTGGAAAAAGACTTT	2522
QY	397	ProProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeu	416	QY	757	LysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuGluAsp	776
DB	1443	CCTCCTCTGTGAACCTCACTGACCCAGACAGATTCACACTCTTTTACCATTTCAGTCTC	1502	DB	2523	AAACTATTTAAAAAATTTTCTCTCTGGAATTAATAACAGATTTTACTTTGGAAGAC	2582
				QY	777	ThrProArgGlnCysArgIleCysGlyLeuAlaMetTyrGluCysArgGluCysTyr	796

Db	2583	ACTCCAGACAGTCCGGATATGTGAGGGCTGCAATGTATGAGTGTAGAGATCTCTAC	2642
Qy	797	AspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGln	816
Db	2643	GACGATCCGGACATCTCAGCTGGAAATAACAGCAGTTTTGTAAAACTGCAACTCAA	2702
Qy	817	ValHisLeuHisProLysArgLeuAenHisLysTyraAsnProValSerLeuProLysAsp	836
Db	2703	GTCCACCTTCATCCGAAAGAGGCTGAATATAAATAAACCCAGTGCATCTCCCAAGAC	2762
Qy	837	LeuProAspTpaAspTpaArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAla	856
Db	2763	TTACCGACTGGAGCTGGAGACACGGCTGCATCCCTTGCAGATATGAGTTATTGCT	2822
Qy	857	ValLeuCysIleGluThrSerHisTyraValAlaPheValLysTyrglyLysAspSer	876
Db	2823	GTCTCTGCATAGAAACAGCCACTATGTGCTTTTGTGAAGTATGGGAAGCAGATTCT	2882
Qy	877	AlaTpaLeuPhePheAspSerMetAlaAspArgAspGlyGlnAsnGlyPheAsnIle	896
Db	2883	GCCTGGCTCTCTTGTACAGACATGGCCGATCGGATGGTGGTGCAGATGGCTTCAACAT	2942
Qy	897	ProGlnValThrProCysProGluValGlyGluTyraLeuLysMetSerLeuGluAspLeu	916
Db	2943	CCTCAAGTCACCCCATGCCAGAGTAGGAGAGTACTTGAAGATGCTCTGGAAGACCTG	3002
Qy	917	HisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTy	936
Db	3003	CATCTCTGGACTCCAGAGAGATCCAAAGCTGTGCAGAGAGCTGTTGTGATGCATAT	3062
Qy	937	MetCysMetTyrglnSerProThrMetSerLeuTyraLys	949
Db	3063	ATGTGATGTACAGAGTCCAAATGAGTTGTACAAA	3101
RESULT 5			
HSA250014			
LOCUS	HSA250014	5371 bp	mRNA linear PRI 16-JUN-2000
DEFINITION	Homo sapiens mRNA for Familial Cylindromatosis cyld gene.		
ACCESSION	AJ250014		
VERSION	AJ250014.1	GI:8250235	
KEYWORDS	cyld gene; Familial Cylindromatosis.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Bignell, G.R., Brown, C., Biggs, P.J., Lakhani, S.R., Jones, C., Hansen, J., Blair, E., Hofmann, B., Siebert, R., Turner, G., Evans, D.G., Schrander-Stumpel, C., Beemer, F.A., Van Den Ouweland, A., Halley, D., Delpech, B., Cleveland, M.G., Leigh, I., Leisti, J., Rasmussen, S., Wallace, M.R., Fenske, C., Banerjee, P., Oiso, N., Chaggar, R., Merrett, S., Leonard, N., Huber, M., Hohl, D., Chapman, P., Burn, J., Swift, S., Smith, A., Ashworth, A. and Stratton, M.R.		
TITLE	Identification of the familial cylindromatosis tumour-suppressor gene		
JOURNAL	Nat. Genet.	25 (2), 160-165	(2000)
MEDLINE	20296617		
PUBMED	10835629		
REFERENCE	2	(bases 1 to 5371)	
AUTHORS	Stratton, M.R.		
TITLE	Direct Submission		
JOURNAL	Submitted (29-NOV-1999) Stratton M.R., Cancer Genetics, Institute of Cancer Research, 15 Cotswold Rd, Sutton., Surrey. SM2 5NG., UNITED KINGDOM		
FEATURES	Location/Qualifiers		
source	1. .5371		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/chromosomes="16"		
	/map="16q12 - q13"		
gene	1. .5371		

CDS	/gene="cyld"		
	392..3262		
	/gene="Cyld"		
	/function="tumour suppressor/recessive oncogene"		
	/note="Familial Cylindromatosis Gene"		
	/codon_start=1		
	/evidence=experimental		
	/protein_id="CA893533.1"		
	/db_xref="GI:8250236"		
	/db_xref="GOA:Q9NQC7"		
	/db_xref="UniProt/TREMBL:Q9NQC7"		
	/translation="MSSGLWSQSKVTSPYWEERIFYLLOECSVTDKOTOKLVKPKVG SIGQYIDRSVGHSRIPSAKGNQIKGLKEQPHAVLFVDEKDVINEKFTLELLA IINCERPSLFNRNRLSKGLQIDVCPVKVQSRSGEKEKPGVVRFGPPLAERTVSG IFFGVLLBEGRGFTDGYVQKQLQFCDEDCGVFVALDKLELIEDDDTALESDDYAG PDGTMQVEIPLPLEINRSVSLKVGTEISGTVIFCDVLPGKESLGYFVGVDMDIGNW DGRFDQVQICSFACVESTILLHNDIIIPALSESVTQRRPPKLAFLMRGVDGKSSSH NKPATGSTDPGNRNSLFTLNGSVSDSQKSKNTWYIDEVADPAKSLTEIS TQDRSSPLOPPVNSLTTFENRPHSLTPKPNNTNGSIGHSPLSLSAQSVMEELN TAPVQSSPPLAMPNGSHGLEVSLAEVKENPEFYGVIRWIGQPPGNEVLGAELED ECAGCTDFTFRGTRYFTCAKALKALFKVLSKCRPSDFASLPQVSNQIERNCSLAFGY LSEVVENTPDKMEKEGLEIMIGKKGIQGHYNSCYLDDSTLFCLEAFSSVLDTVLLRP KEKNDVEYSETEOLLRTIVNPLRYGVYCATKIMKRLKLEKVEAGSGTSEKDP EERFNLTHPHILRVEPLIKRSAGQKVDQCYFQIEMEKKEKVGPTIQQLLEWSFIN SNLKFAEASCLIIQMPRFGKDFLKFIFPSLELNITDLEDTPRCRCICGGLAMYE CRECYDDPDISAGIKQFCNTQVHLHPKRLNHNKYNVPSLPKDLDPDWRHRCIPC QNMELFAVLCIETSHYVAFYKGDSDSAWLFFDSMDRQGGNGFNIPQVTPCPPEVGE YLKMSLEDLSLDSRRIOGCARLLCDAYMCMYQSPMTSLYK"		
ORIGIN			
Alignment Scores:			
Pred. No.:	0	Length:	5371
Score:	4971.50	Matches:	948
Percent Similarity:	99.16%	Conservative:	0
Best Local Similarity:	99.16%	Mismatches:	1
Query Match:	98.76%	Indels:	7
DB:	9	Gaps:	5
US-09-671-687A-3 (1-949) x HSA250014 (1-5371)			
Qy	1	MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTyTrpGluGluArgIle	20
Db	392	ATGATGTTGAGGCTTATGGAGCCCAAGAAAAAGTCACTTCCCTACTGGGAAGCGGAT	451
Qy	21	PheTyraLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLeuLys	40
Db	452	TTTACTTCTCTTCAAGATCGCGTTACAGACAAACACACAAAGCTCTTAAA	511
Qy	41	ValProLysGlySerIleGlnTyraIleGlnAspArgSerValGlyHisSerArgIle	60
Db	512	GTACCGAAGGGAAGTATAGCACAGTATATTCAAGATCGTCTCTGTGGGCGATTCAAGGAT	571
Qy	61	ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla	80
Db	572	CCTTCTGCAAAAGGCAAGAAAAATCAGATTGGATTAAAAATTTAGAGCAACCTCATGCA	631
Qy	81	ValLeuPheValAspGlu---AspValValGluLeuAsnGluLysPheThrGluLeuLeu	99
Db	632	GTTCTCTTTTGTGATGAAAAGGATGTTGTAGAGATAAATGAAAAGTTTACAGAGTTACT	691
Qy	100	LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer	119
Db	692	TTGGCAATTACCAATTGTGAGGAGAGGTTTCAGCGCTGTTTAAAAACAGAAACAGACTAAGT	751
Qy	120	LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu	139
Db	752	AAAGGCTCCAAATAGACGTGGGCTGTCTGTGTAAGTACAGCTAGATCTCGGGAAGAA	811
Qy	140	LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly	159
Db	812	AAATTTCTGGAGTTGTACGCTTCAGAGGACCCCTGTTAGCAGAGACAGTCTCCCGA	871
Qy	160	IlePhePheGlyValGluLeuLeuGluGluGlyArgGlyGlnGlyPheThrAspGlyVal	179

QY 894 PheAsnIleProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeu 913
 Db 3092 TTCAACATTCCTCAAGTCACCCCATGCCAAGAGTAGGAGACTTTGAAGATGTCCTG 3151
 QY 914 GluAspLeuHisSerLeuAspSerArgGileGlnGlyCysAlaArgArgLeuLysCys 933
 Db 3152 GAAGACCTGCATTCCTTGGACTCCAGAGAGATCAAGGCTGTGCACGAAGACTGCTTTGT 3211
 QY 934 AspAlaTyrMetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
 Db 3212 CATGCATATATGTCATGTACCAAGAGTCCCAACAATGAGTTGTACAAA 3259
 RESULT 6
 CQ834226
 LOCUS CQ834226 3311 bp DNA linear PAT 29-JUL-2004
 DEFINITION Sequence 97 from Patent WO2004058805.
 ACCESSION CQ834226
 VERSION CQ834226.1 GI:50833763
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 AUTHORS Matsuda, A. and Yoneta, S.
 TITLE T cell activating gene
 JOURNAL Patent: WO 2004058805-A 97 15-JUL-2004; -late
 Asahi Kasei Pharma Corporation (JP)
 FEATURES
 source
 1. 3311
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 243.. 3113
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAH05328.1"
 /db_xref="GI:50833764"

CDS

/translation="MSSGLWSQEKVTSPYWEERI FYLLQECVSDTKQKLLKVPKG
 SIGQYIDRSVGHRI PSAGKKNQIGLILQPHAVL FDEKDVINEKE TELLLA
 INCSERFLFKNRNRLSGQLIDVCPVKVQLRSEEPFGVRRFRGPLAERTVSG
 IFPGVELLEEGQGQDGVYQKQLFQCDGCEVVALDKLELIEDDDTALSDYAG
 PGDTQVPELLEINRSVLKGETTESGTVPFCDLPKGSLSGYEVGDMNPINW
 DGRFGVQLCSFACVETILLINDIIIPALSESVOERPPKLFAMSRGVGDKSSSH
 NKPKATGSDPDGNRRNSELVYTLNGSSVDSQPSKNTWYIDEVAEPKSLTEIS
 TPDFSSPLQPPVNSLTENRFHSLPSLTMPNTNGSIGHSPLSLSAQSVWSELN
 TAPQESPLAMPNGSHGLEVSLAEVRENPFYGVIRWIGQPPGLNEVLGLELED
 ECAGCTDGRFTRYPTCAKLFVKLJSCRPDSRFSALQPVNSQIERCNSLAFGY
 LSEVVEENTPPRMEKEGLIMIGKKGIQHYNSCYLDSLTFLCFAPSSVLDVLLRP
 KEKNVVEYSETQELARTVINPLRYGVYCATKIMKLLKLEKVEAAGFTSEKDP
 BEFLNLTFHHLIRVEPLKIRSAGQVQDCYFQIEMEKNEKVGPTIOQLLEWSPIN
 SNLKFAEARSCLIIQMPREGKDFPKKIFPSLELNITLLEDTPRQCRICGLLWYE
 CRECDPDISAGIKQFCKNTQNTQHLHPKRLNHNYPVSLPDLPDLPDWRHGCIPPC
 QNMFELVICIETSHYVAFVYKGDSDSAWLFDSMDRGGQNGENIPQVTPCPEVE
 YLKMSLEDLHSLDSRRIQGCARRLLCDAYMCMYQSPMSLYK"

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 3311
 Score: 4968.50 Matches: 947
 Percent Similarity: 99.16% Conservative: 1
 Best Local Similarity: 99.06% Mismatches: 7
 Query Match: 98.70% Indels: 1
 DB: 6 Gaps: 5

US-09-671-687A-3 (1-949) x CQ834226 (1-3311)

QY 1 MetSerSerGlyLeuTyrSerGlnGluLysValThrSerProTyrTrpGluGluArgGile 20
 Db 243 ATGAGTTCAGGTTATGGAGCCCAAGAAAAGTCACTTCCACCTACTGGGAAGCGGATT 302
 QY 21 PheTyrLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLys 40

Db 303 TTTTACTTGTCTTCAAGAAATGAGCGTTTACAGACAAACAAACAAAGCTCTCTTAAA 362
 QY 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgile 60
 Db 363 GTACCGAAGGAAAGTATAGACACTATATTCAAGATCGTCTGTGGGCAATCAAGGATT 422
 QY 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla 80
 Db 423 CCTTCTGCAAAAGCAAGAAAATCAGATTGGATTAATAAATTTAGAGCAACCTCATGCA 482
 QY 81 ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu 99
 Db 483 GTTCTCTTGTGATGAAAAGGATGTTGTAGAGATAAATGAAAAGTTTACAGAGTTACTT 542
 QY 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
 Db 543 TTGGCAATTACCAATTGTGAGGAGAGGTTGAGCTGTTTAAACACAGAAACAGACTAAGT 602
 QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
 Db 603 AAAGGCTCCAAAATAGACGTGGGCTGCTGTGAAAAGTACAGCTGAGATCTGGGGAAGAA 662
 QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
 Db 663 AAATTTCTGAGGTTGTACGCTTCAGAGGACCCCTGTTAGCAGAGAGACAGTCTCCGGA 722
 QY 160 IlePhePheGlyValGluLeuLeuGluGluGlyArgGlyGlnGlyPheThrAspGlyVal 179
 Db 723 ATATTCTTTGAGGTTGAATTTGCTGGAAGAGGCTGCTGTCGAGGTTTCACTGACGGGTG 782
 QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
 Db 783 TACCAAGGGAACACAGCTTTTTCAGTGTGATGAAGATTGTGCGGTGTTGTTGTCATTGGAC 842
 QY 199 LysLeuGluLeuLeuGluAspAspThrAlaLeuLeuGluSerAspTyrAlaGlyProGly 218
 Db 843 AAGCTAGAACTCATAGAAGATGATGACACTGCTGTTGAAAGTGAATTAACGAGGCTCTGGG 902
 QY 219 AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly 238
 Db 903 GACAAATGAGGTCGAACTTCTCTCTTGGAAATAAACTCCAGAGTTCTTTGAAGTT 962
 QY 239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
 Db 963 GGAGAAACAATAGAACTCTGGAACAGTTATCTCTGTGATGTTTGTCCAGGAAAGAAAGC 1022
 QY 259 LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe 278
 Db 1023 TTAGGATATTTTGTGTTGGTGGACATGATAACCTATTGGCAACTGGGATGGAAGATT 1082
 QY 279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuHisIleAsn 297
 Db 1083 GATGGAGTGCAGCTTTGTAGTTTTCGCTGTGTGAAAAGTACAAATCTTATTGCACATCAAT 1142
 QY 298 AspIleIlePro-----GluSerValThrGlnGluArgArgProProLysLeuAla 314
 Db 1143 GATATCATCCAGCTTTTATCAGAGAGTGTGACGAGGAAAGAGGCGCTCCCAACTTCC 1202
 QY 315 PheMetSerArgGlyValGlyAspLysGlySerSerSerHisAsnLysProLysAlaThr 334
 Db 1203 TTTATGTCAAGAGGTGTGGGACAAAGGTTTCCAGCTCATATAAACCACCAAGGCTACA 1262
 QY 335 GlySerThrSerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrIleuAsnGly 353
 Db 1263 GGATCTACCTCAGACCCCTGGAAATAGAAACAGATCTGAAATTAATTTATACCTTAAATGGG 1322
 QY 354 SerSerValAspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGluVal 373
 Db 1323 TCTTCTGTTGACTCACAAACCAATCCAAATCAAAATAATATGATGATCATGATGAGTT 1382
 QY 374 AlaGluAspProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerPro 393

Db 1383 GCAGAGACCCCTGCATAATCTCTTACAGAGATATCTACAGACTTTTGACCGTTCTTCCACCA 1442
Qy 394 ProLeuGlnProProValLeuSerLeuThrThrGluAsnArgPheHisSerLeuPro 413
Db 1443 CCACCTCCAGCCTCCTCTGTGAACCTCACTGACCCAGAGAACAGATTCACACTCTTTACCA 1502
Qy 414 PheSerLeuThrLysMetProHsnThrAsnGlySerIleGlyHisSerProLeuSerLeu 433
Db 1503 TTCAGTCTCACCAGATGCCCCAATCAATGGAAGTATTGGCCACAGTCCACTTTCTCTG 1562
Qy 434 SerAlaGlnSerValMetGluLeuLeuAsnThrAlaProValGlnGlnSerProProLeu 453
Db 1563 TCAGCCCACTCTGTANTGAGAGCTTAACACTGACCCGCTCCAGAGAGTCCACCCCTTG 1622
Qy 454 AlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGlu 473
Db 1623 GCCATGCCCTCTGGGAACCTCACATGCTTAGAAGTGGGCTCATTTGGCTGAAGTTAAGGAG 1682
Qy 474 AsnProProPheThrGlyValIleArgTrrPilleGlyGlnProProGlyLeuAsnGluVal 493
Db 1683 AACCCCTCTCTTATGGGGTAATCCGTTGGATCGGTGAGCCACAGGACTGAATGAAGTG 1742
Qy 494 LeuAlaGlyLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGly 513
Db 1743 CTCGCTGGAGCTGAACCTGGAAGATGAGTGTGAGGCTGTACGGATGGAACCTTCAGAGGC 1802
Qy 514 ThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArg 533
Db 1803 ACTCGGTATTTCACCTGTGCTGAGAGAGGCGCTGTTGTGAACTGAGAGCTGCAGG 1862
Qy 534 ProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSer 553
Db 1863 CCTGACTCTAGGTTTCATCTAGTGCAGCCGCTTCCAAATCAGATTGAGCGCTGTAACCTCT 1922
Qy 554 LeuAlaPheGlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGlu 573
Db 1923 TTAGCATTTGGAGGCTACTTAAAGTAGAGTAGTAGAAGAAATACTCCACCAAAATGGAA 1982
Qy 574 LysGluGlyLeuGluIleMetIleGlyLysLysLysGlyIleGlnGlyHisTyrAsnSer 593
Db 1983 AAAGAGGCTTGGAGATATGATTGGAGAGAGAGAGGATCCAGGTCATTTACATTTCT 2042
Qy 594 CysTyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrVal 613
Db 2043 TGTACTTAGACTCAACCTATTCTCTATTGCTTTTGTAGTCTGTTCTGGACACTGTG 2102
Qy 614 LeuLeuArgProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeu 633
Db 2103 TTACTTAGACCCCAAGAGAAAGAACGATGTAGATATTTATAGTGAACCCCAAGAGCTACTG 2162
Qy 634 ArgThrGluIleValLeuAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMet 653
Db 2163 AGACAGAAATTTGTAATCTCTGAGAAATATATGGATATGTGTGTCACAAAATATG 2222
Qy 654 LysLeuArgLysIleLeuGluLysValGluAlaSerGlyPheThrSerGluGluLys 673
Db 2223 AAACCTGAAGAAATACTTGAAGGTGGAGGCTGCATCAGGATTTACCTCTGAGAGAAA 2282
Qy 674 AspProGluGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeu 693
Db 2283 GATCCTTGAGGAATTTTGAATATCTGTTTCATCATATTTTAAAGGGTAGAACCTTTGCTA 2342
Qy 694 LysIleArgSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGlu 713
Db 2343 AAATAAGATCAGCAGGTCAAAAGGTACAGATTTGTACTTCTATCAAAATTTTATGGAA 2402
Qy 714 LysAsnGluLysValGlyValProThrIleGlnGlnLeuLeuGluTrrPheIleAsn 733
Db 2403 AAAAATGAGAAAGTTCGCGTTCACCAATTCAGAGTGTGTAGATGGCTTTTATCAAC 2462
Qy 734 SerAsnLeuLysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGly 753
Db 2463 AGTAACCTGAAATTTGAGAGGACCATCATGTCTGATTATTACAGATGCCTCGATTGGA 2522

Qy 754 LysAspPheLysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeu 773
Db 2523 AAAGACTTTAAACTATTATAAAAAATTTTCTCTCTGGAATTAATAATAACAGATTTA 2582
Qy 774 LeuGluAspThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArg 793
Db 2583 CTTGAAGACACTCCCAAGACAGTCCCGATATGAGAGGGCTTGCATGTATGATGTAGTA 2642
Qy 794 GluCysTyrAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCys 813
Db 2643 GAATGCTACGACGATCCGACATCTCAGCTGGAAAAATCAAGCAGTTTGTAAAACTGC 2702
Qy 814 AsnThrGlnValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeu 833
Db 2703 AACACTCAAGTCCACTCTCATCCGAAGAGGCTGAATCATATAATAACCCAGTGTCACTT 2762
Qy 834 ProLysAspLeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGlu 853
Db 2763 CCAAAAGACTTACCCGACTGGGACTGGAGACAGCTGCATCCCTGCCAGAAATATGAG 2822
Qy 854 LeuPheAlaValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLys 873
Db 2823 TTATTGCTGTTCTCTGATAGAAACAAGCCACTATGTTCTTTGTGAAGTATGGGAAG 2882
Qy 874 AspAspSerAlaThrLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGly 893
Db 2883 GAGGATTCCTCGCTGGCTCTCTTTGACAGCATGCCGATCGGATGTGTGTCAGATGGC 2942
Qy 894 PheAsnIleProGlnValThrProCysProGluValGlyGlyTyrLeuLysMetSerLeu 913
Db 2943 TTCAACATTCCTCAAGTCAACCCCATGCCAGAAAGTAGGAGACTTTGAAGATGCTCTG 3002
Qy 914 GluAspLeuHisSerLeuAspSerArgIleGlnGlyCysAlaArgArgLeuLysCys 933
Db 3003 GAAGACTGCTATTCCTTGGACTCCAGAGAAATCAAGGCTGTGCACGAAGACTCTTGT 3062
Qy 934 AspAlaTyrMetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
Db 3063 GATGCATATATGTGCATGTACACAGAGTCCAAACATGAGTTGTATACAA 3110

RESULT 7

LOCUS CQ719792 3480 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 5726 from Patent WO02068579.
ACCESSION CQ719792
VERSION CQ719792.1 GI:42280649
KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1

AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.

TITLE Kits, such as nucleic acid arrays, comprising a majority of

humanexons or transcripts, for detecting expression and other uses

JOURNAL thereof WO 02068579-A 5726 06-SEP-2002;

PE Corporation (NY) (US)

FEATURES

source 1.3480

Location/Qualifiers

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 3480

Score: 4962.00 Matches: 947

Percent Similarity: 99.27% Conservatives: 0

Best Local Similarity: 99.27% Mismatches: 2

Query Match: 98.57% Indels: 5

DB: 6 Gaps: 4

```
US-09-671-687A-3 (1-949) x CQ719792 (1-3480)
QY 1 MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluGluArgIle 20
Db 235 ATGAGTTTCAGGCTTATGAGGCCAAGAAAAGTCACTTCACCTTCTGCGGAAGACGGATT 294
QY 21 PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLys 40
Db 295 TTITACTTCTCTTCAAGAAATGACGGTTTACAGACAAACAAACAAAGCTCTCTAAA 354
QY 41 ValProLysGlySerIleGlyLysGlnTrpIleGlnAspArgSerValGlyHisSerArgIle 60
Db 355 GTACCGAAGGAGATAGACAGATATATTCAGATCGTCTCTGCGGCAATTCAGGATT 414
QY 61 ProSerAlaIleGlyLysLysGlnIleGlyLeuLysIleLeuGluGlnProHisAla 80
Db 415 CCTTCTGCAAAAGGCCAAGAAAATCAGATTGGATTAAATAATCTAGAGCAACCTCATGCA 474
QY 81 ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu 99
Db 475 GTTCTCTTTGTTGATGAAAGGATGTTGTAGAGATAAATGAAAGTTTCAGAGATTACTT 534
QY 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
Db 535 TTGGCAATTTACCAATTGTGAGAGAGGTTGAGCTGTTTAAATAACAGAACAGACTAAGT 594
QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
Db 595 AAAGGCTCCAAATAGACGCTGGCTCTCTGTGAAAGTACAGCTGAGATCTGGGGAAGAA 654
QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
Db 655 AAATTTCTCTGGAGTTGTACGCTTTCAGAGACCCCTGTAGACAGAGGACAGCTCTCCGA 714
QY 160 IlePhePheGlyValGluLeuLeuGluGlyArgGlyGlnGlyPheThrAspGlyVal 179
Db 715 ATATTCTTTGGAGTTGAAATGCTGGAAGAGGTCGTGGTCAAGGTTTCACTGACGGGGTG 774
QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
Db 775 TACCAGGGAACACGCTTTTTCAGTGTGATGAAGATTGTGGCGTGTGTTGTGCAATGGAC 834
QY 199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
Db 835 AAGCTAGAACTCATAGAACATGATGACATGCAATTCGAAAGTATTACGACGCTCTGGG 894
QY 219 AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly 238
Db 895 GACACAATGCAGGTCGAACCTCTCTCTTTGGAAATAAATCCACAGATTTCTTTGAAAGTT 954
QY 239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
Db 955 GGAGAAACAAATAGATCTGGAACAGTATATCTGTGATGTTTTCAGAGAAAGAAAGC 1014
QY 259 LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe 278
Db 1015 TTACGATATTTTGTGTTGGTGTGACATGGATAACCTATTGGCAACTGGGATGGAAGATT 1074
QY 279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
Db 1075 GATGAGTGCAGCTTTGTAGTTTGTGCTGTGTTGAAAGTACAAATCTTATTGCAATCAAT 1134
QY 298 AspIleIleProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMetSer 317
Db 1135 GATATCATCCAGAGAGTGTGACGAGAAAGGAGGCTCCCAAACTTGCCTTTATGTCA 1194
QY 318 ArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThrGlySerThr 337
Db 1195 AGAGGTGTTGGGACAAAGGTTTCATCCAGTCATATAATAACCAAGGCTACAGATCTACC 1254
QY 338 SerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVal 356
Db 1255 TCAGACCTCGAAATAGAAACAGATCTGAATATATTTTATACCTTAAATGGGTCTTCTGT 1314
QY 357 AspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGlu-ValAlaGluAs 376
Db 1315 GACTCAACACCAATCAATCAAAATCAAAATACATGTTGATGATGAGGTTGCAGAAAGA 1374
QY 376 pProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerProProLeuGlu 396
Db 1375 CCTGCAAAATCTCTTACAGAGATATCTACAGACTTTGACCGTCTTCCACCACACTCCA 1434
QY 396 nProProProValAsnSerLeuThrGluAsnArgPheHisSerLeuProPheSerLe 416
Db 1435 GCCTCTCTCTGTGAACCTCACTGACCCACGAGAACAGATTCCACTCTTTTACCATTCAGTCT 1494
QY 416 uThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGlu 436
Db 1495 CACCAAGATGCCCAATACCAATGGAAGTATTGGCCACAGTCCACTTTCTCTGTGAGCCCA 1554
QY 436 nSerValMetGluLeuLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPr 456
Db 1555 GTCTGTAATGGAAGAGCTAAACACTGCACCCGTCCAAGAGAGTCCACCTTTGGCCATGCC 1614
QY 456 oProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPr 476
Db 1615 TCCTGGGAACCTACATGCTCTAGAGTGGGCTCATTTGGCTGAAGTTAAGGAGAACCTCTCC 1674
QY 476 oPheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGlu 496
Db 1675 TTTCTATGGGTAAATCGTTGGATCGGTACGCCACACGAGTGAATGAAGTCTCGCTGG 1734
QY 496 yLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTy 516
Db 1735 ACTGSAACCTGGAAGAGTGAAGTGCAGGCTGTACGGATGGAACTTCAGAGGACCTCGGTA 1794
QY 516 rPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSe 536
Db 1795 TTTCACTGTGCCCTGAAGAGGCGCTGTTGTGAAACTGAAAGAGTGCAGGCTGACTC 1854
QY 536 rArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPh 556
Db 1855 TAGTTTTCATCATTTGAGCCGGTTTCCATTCAGATTGAGCGTGAACCTCTTTAGCATT 1914
QY 556 eGlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluGlu 576
Db 1915 TGGAGGCTACTTAAGTGAAGTAGTAGAAGAAATACTCCACCAAAATATGGAAGAAAGG 1974
QY 576 yLeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLe 596
Db 1975 CTTGGAGATAATGATTGGGAAGAAAGAGGATCCAGGGTCATTACAAATCTTGTACTT 2034
QY 596 uAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuAr 616
Db 2035 AGACTCAACCTTATTTCTGTTATTTGCTTTTAGTTCTGTTCTGACACTGTGTTACTTAG 2094
QY 616 gProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu 636
Db 2095 ACCCAAGAAAGAACGATGAGATAATATTAGTGAACCCAGAGCTACTTGAGGACAGA 2154
QY 636 ulleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuAr 656
Db 2155 AATTGTTAACTCTCTGAGAAATATGAGATATGTGTGTCACCAAAATATATGAAACTGAG 2214
QY 656 gLysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGlu 676
Db 2215 GAAATATCTGAAAGGTTGAGGCTGCATCAGGATTTTACCTCTGAGAGAAAGATCTCTGA 2274
QY 676 uGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleAr 696
Db 2275 GGAATCTTGAATATCTGTTTTCATCATATATTTAAAGGGTAGAACCTTTGCTAAATAAG 2334
QY 696 gSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu 716
Db 2335 ATCAGCAGGTCAAAGGTTACAGATTGTTACTTCTATCAATTTTTTATGGAATAAATGA 2394
```


QY	716	uLysValGlyValProThrIleGlnGlnLeuLeuGluTripSerPheIleAsnSerAsnLe	736
DB	2395	GAAGAATGGCGTTCCCAACAATTGAGAGTGTGTAGAAATGGTCTTTTATCAACAGTAACCT	2454
QY	736	uLysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPh	756
DB	2455	GAATAATTTGACAGAGCCACCATGCTCGATTTATTTCAGATGCCCTCGATTTGCGAAAAGACTT	2514
QY	756	eLysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAs	776
DB	2515	TAAACTATTATAAAAAATTTTCTCTCTCGGAATTAAATATAACAGATTTACTTGAAGA	2574
QY	776	pThrProArgGlnCysArgIleCysGlyLeuAlaMetTyrGluCysArgGluCysty	796
DB	2575	CACCTCCAGACAGTGC CGGATATGTGAGGCGCTTGCAATGTATGAGTGTAGAGAATGCTA	2634
QY	796	rAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGl	816
DB	2635	CGACGATCCGAGACATCTCAGCTGGAAAATAACAGCAGTTTTTGTAAAAACCTGCAACACTCA	2694
QY	816	nValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAs	836
DB	2695	AGTCCACCTTCATCCGAAGAGCGTGAACTAATATATAACCCAGTGTCACTTCCCAAGA	2754
QY	836	pLeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAl	856
DB	2755	CTTACCGACTGGGACTGGACACCGGTGCATCCCTTGCAGAAATATGGAGTTATTTCG	2814
QY	856	aValLeuCystIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSe	876
DB	2815	TGTTCTCTGCATAGAAACAACCCACTATGTTGCTTTTGTGAAGTATGGGAAGACGATTC	2874
QY	876	rAlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIl	896
DB	2875	TGCTTGGCTCTTCTTGACACATGGCCGATCGGGATGGTGGTCAGAAATGGCTTCAACAT	2934
QY	896	eProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLe	916
DB	2935	TCCTCAAGTCACTCCCATGCCAGAAAGTAGGAGAGTACTTGAAGATGTCTCTGGAAGACCT	2994
QY	916	uHisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTy	936
DB	2995	GCATTCCTTGGACTCCAGAGAAATCCAAAGCTGTGCAAGAAAGCTGCTTGTGTATGATCA	3054
QY	936	rMetCysMetTyrGlnSerProThrMetSerLeuTyrLys	949
DB	3055	TATGTGATGTATCAGAGTCCAAACAAATGAGTTGTACAAA	3094

RESULT 8	BD231207	BD231207	4527 bp	DNA
LOCUS	BD231207	Human cytoskeleton associated proteins.		linear
DEFINITION	BD231207			
ACCESSION	BD231207.1	GI:33040977		
VERSION	JP 2002526076-A/9.			
KEYWORDS	Human sapiens (human)			
SOURCE				

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 4527)

AUTHORS
Lal, P., Tang, T. Y., Yue, H., Hillman, J. L., Bandman, O., Corley, N. C., Guegler, K. J., Patterson, C., Azimzai, Y. and Baughn, M. R.

TITLE Human cytoskeleton associated proteins
JOURNAL Patent: JP 2002526076-A 9 20-AUG-2002; -- late

COMMENT	OS	Homo sapiens (human)	INCYTE PHARMACEUTICALS INC

PN	JP 2002536076-A/9
PD	20-AUG-2002
SZ	17 JUL 2002 TD 0000ZTTT02F4

PF 17-SEP-1999 JP 2000574254
PR 18-SEP-1998 US 60/172226, 27-APR-1999 US 60/131321 PI
PBRETTI LAI TOM V TANG HENDY VIE JENNIFER I. HILLMAN OJCA BT

FREEDT LEE, TOM T LEE, HENRY TOE, VERNITTEN E HATTERTY, OUCH I I
BANDMAN,
BANDMAN,

PI	NEIL C CORLEY, KARL J GUEGLER, CHANDRA PATTERSON, YALDA AZIMZAI,
PI	MARIAH R BAUGHN
PC	C12N15/09,A61K38/00,A61K45/00,A61P1/16,A61P3/00,A61P3/06 PC
PC	A61P3/10,A61P5/00,
PC	A61P5/38,A61P7/00,A61P7/06,A61P9/00,A61P9/10,A61P11/06 PC
PC	A61P13/00,A61P13/12,
PC	A61P17/06,A61P19/02,A61P19/10,A61P21/04,A61P25/00,A61P25/00,
PC	A61P25/04,
PC	A61P25/14,A61P25/16,A61P25/28,A61P29/00,A61P31/04,A61P31/12,
PC	A61P31/18,
PC	A61P33/00,A61P35/00,A61P35/02,A61P37/00,A61P43/00,C07K14/47,
PC	C07K16/18,
PC	C12N1/15,C12N1/19,C12N1/21,C12N5/10,C12P21/02,C12P21/08 PC
PC	C12Q1/68,G01N33/15,
PC	G01N33/50,G01N33/53,G01N33/53,G01N33/566,C12N15/00,C12N5/00,
PC	A61K37/02
CC	Incyte ID No: 2363327
FT	Key
FT	source
FT	Location/Qualifiers
FT	1..4527
FT	/organism="Homo sapiens (human)"
FT	Location/Qualifiers
FT	1..4527
FT	/organism="Homo sapiens"
FT	/mol_type="genomic DNA"
FT	/db_xref="taxon:9606"
ORIGIN	
Alignment Scores:	
Pred. No.:	0
Score:	4942.00
Percent Similarity:	99.37%
Best Local Similarity:	99.37%
Query Match:	98.17%
DB:	6
US-09-671-687A-3 (1-949) x BD231207 (1-4527)	
Qy	4 GlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluGluArgIlePheTyrIleu 23
Db	270 GGCTTATGGAGCCAAAGAAAAGTCACTTCACCCCTACTGGGAAGAGCGGATTTTTTACTTGT 329
Qy	24 LeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLeuLysValProLys 43
Db	330 CTTCTTCAAGAATCGACGGTTCACAGCAACAAACACAAAAGTCTCTTAAGTACCGAA- 388
Qy	44 GlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIleProSerAla 63
Db	389 GGAAGTATAGCAGCATGATATTTCAGATCGTTCTGTGGGCAATTCAGGATTCCTCTGCA 448
Qy	64 LysGlyLysLysAsnGlnIleGlyLeuLysValIleLeuGlnProHisAlaValLeuPhe 83
Db	449 AAAGCGCAAGAAAATCAGATTGGATTAAAAATTCAGAGCAACCTCATGCAAGTTCTCTTT 508
Qy	84 ValAspGlu--AspValValGluIleAsnGluLysPheThrGluLeuLeuAlaIle 102
Db	509 GTTGATGAAAAGGATGTTGTAGATTAATGAAAAGTTCCAGAGTTACTTTTGGCAATT 568
Qy	103 ThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSerLysGlyLeu 122
Db	569 ACCAATTGTGAGGAGAGTTTCAGGCCTGTTAAAAACAGAAACAGACTAAGTAAAGGCGCTC 628
Qy	123 GlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGluLysPhePro 142
Db	629 CAATATAGCGTGGCTGTCTGTGAAAGTACAGCTGAGATCTGGGGAAGAAAAATTTCTT 688
Qy	143 GlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGlyIlePhePhe 162
Db	689 GGAGTTGTACCTTCAGAGGACCCCTGTTCAGAGAGGACAGCTCCGGAATATTTCTT 748
Qy	163 GlyValGluLeuLeuGluGluGlyArgGlyGlnGlyPheThrAspGlyValTyrGlnGly 182
Db	749 GGAGTTGAATTCCTGGAAAGAGTCTGGTCAAGGTTTCACTGACGGGGTGTACCAAGG 808

QY 183 LysGlnLeuPheGlnCysAspGluAaspCysGly---PheValAlaLeuAaspLysLeuGlu 201
DB 809 AAACAGCTTTTTCAGTGTGATGAAGATTGTGGCGTGTGTTGTCATTGGACAAGCTAGAAA 868
QY 202 LeuIleGluAAspAspThrAlaLeuGluSerAspTyrAlaGlyProGlyAspThrMet 221
DB 869 CTCATAGAGATGATGACACTGCTATTGGAAAGTGATTACGCAGGTCTCTGGGGACACAATG 928
QY 222 GlnValGluLeuProProLeuGluIleAasnSerArgValSerLeuLysGlyGluThr 241
DB 929 CAGGTGCAACTCTCTCTTGGAAATAAACTCCAGAGATTCTTTGAAGGTTGGAGAAACA 988
QY 242 IleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSerLeuGlyTyr 261
DB 989 ATAGAACTGGAAACAGTTATTTCTGTGATGTTTTCAGGAAAGAAAGCTTAGGATAT 1048
QY 262 PheValGlyValAspMetAspAenProIleGlyAsnTyrAspGlyArgPheAspGlyVal 281
DB 1049 TTTGTTGGTGTGGACATGGATTAACCTATTGGCACTGGATGGAGATTGTGATGGAGTG 1108
QY 282 ---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAasnAspIleIle 300
DB 1109 CAGCTTTGTAGTTTTCGGTGTGTTGAAAGTACAAATCTATTGCACATCAATGATATCATC 1168
QY 301 ProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMetSerArgGlyVal 320
DB 1169 CCAGAGAGTGTGACGCGAGAAAGAGGCGCTCCCAAACTTGGCTTTTATGTCAAGAGGTGTT 1228
QY 321 GlyAspLysGlySerSerHisAasnLysProLysAlaThrGlySerThrSerAspPro 340
DB 1229 GGGGCAAAAGGTTCACTCCAGTCATAATAAACCAAGGCTACAGGATCTACCTCAGACCCCT 1288
QY 341 GlyAsnArg---ArgSerGluLeuPheTyrThrLeuAasnGlySerSerValAspSerGln 359
DB 1289 GGAATAGAAACAGACTGAAATATTATTTATACCTTAAATGGGTCTCTGTGTGACTCACAA 1348
QY 360 ProGlnSerLysSerLysAsnThrTyrIleAspGluValAlaGluAspProAlaLys 379
DB 1349 CCACAATCCAAATCAAAAAATACATGGTACATTTGATGAAGTTGCAGAAAGACCCCTGCAAAA 1408
QY 380 SerLeuThrGluIleSerThrAspPheAspArgSerSerProLeuGlnProProPro 399
DB 1409 TCTCTTACAGAGATATCTACAGACTTTGACCGTCTCTCCACCACCTCCAGCCTCTCTCCT 1468
QY 400 ValAsnSerLeuThrThrGluAasnArgPheHisSerLeuProPheSerLeuThrLysMet 419
DB 1469 GTGAACCTACTGACCCACCGAGAAACAGATTCCATCTTTACCATTCAGTCTCACCAAGATG 1528
QY 420 ProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGlnSerValMet 439
DB 1529 CCCAATACCAATGGAAGTATTGGCCACAGTCCACTTTCTGTGACGCCAGCTCTGTAATG 1588
QY 440 GluGluLeuAasnThrAlaProValGlnGluSerProProLeuAlaMetProProGlyAsn 459
DB 1589 GAAGAGCTAAACACTGCACCCGTCCAAGAGAGTCCACCCCTTGGCCATGCTCTCTGGGAAC 1648
QY 460 SerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAasnProProPheTyrGly 479
DB 1649 TCACATGGTCTAGAAGTGGGCTCATTTGGCTGAAGTTAAGGAGAACCTCTCTTCTATGGG 1708
QY 480 ValIleArgTyrIleGlyGlnProProGlyLeuAasnGluValLeuAlaGlyLeuGluLeu 499
DB 1709 GTAATCCGTTGGATTCGGTCAGCCACCAAGACTGAATGAAGTGTCTCGCTGGACGTGGAACTG 1768
QY 500 GluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyrPheThrCys 519
DB 1769 GAAGATGAGTGTGAGGCTGTACGGATGGAACCTTCAGAGGCCTCGGTATTTTACCTGT 1828
QY 520 AlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSerArgPheAla 539
DB 1829 GCCCTGAAGAGGCGCTGTTGTGAAACTGAAGAGCTGCAGGCTGACTCTAGGTTTGCA 1888
QY 540 SerLeuGlnProValSerAasnGlnIleGluArgCysAasnSerLeuAlaPheGlyGlyTyr 559

DB 1889 TCATTGACGCGGTTTCCAATCAGATTGACGCTGTAACTCTTTAGCAATTTGGAGGCTAC 1948
QY 560 LeuSerGluValValGluGluAasnThrProProLysMetGluLysGluGlyLeuGluIle 579
DB 1949 TTAAGTGAAGTAGTAGAAGAAATACTCCACCAAAATGGAAAAGAGGCTTCGAGATA 2008
QY 580 MetIleGlyLysLysLysGlyIleGlnGlyHisTyrAasnSerCysTyrLeuAaspSerThr 599
DB 2009 ATGATTGGGAAGAAAGAGCATCCAGGTCATTACAAATCTCTGTTACTTAGACTCAACC 2068
QY 600 LeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuAtpProLysGlu 619
DB 2069 TTATTCTGCTATTGCTTTTAGTCTGTGTCGACACTGTGTACTTAGACCCCAAGAA 2128
QY 620 LysAasnAspValGluTyrTyrSerGluThrGlnGluLeuLeuAtpThrGluIleValAasn 639
DB 2129 AAGACGATGTAGANATATTATAGTGAACCCAGAGCTACTGAGACAGANAATTTGTTAAT 2188
QY 640 ProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuAtpLysIleLeu 659
DB 2189 CCTCTGAGATATATGATATGTGTGCCACAAAAATTTATGAAACTGAGGAAAAATACTT 2248
QY 660 GluLysValGluAlaSerGlyPheThrSerGluGluLysAspProGluGluPheLeu 679
DB 2249 GAAAAGGTGGAGGCTGCATCAGGATTTACCTCTGAAGAAAAGATCTCTGAGANAATTTCTTG 2308
QY 680 AsnIleLeuPheHisIleLeuAtpValGluProLeuLeuLysIleAtpSerAlaGly 699
DB 2309 AATATTCTGTTTCATCANATTTTAAAGGTAGAACCTTTGCTPAAAAATTAAGATCAGCAGT 2368
QY 700 GlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAasnGluLysValGly 719
DB 2369 CAAAAGGTACAAGATTGTTACTTCTATCAAAATTTTATGGAATAAATGAGAAAGTTGGC 2428
QY 720 ValProThrIleGlnGlnLeuLeuGluTyrPheIleAasnSerAasnLeuLysPheAla 739
DB 2429 GTTCCACCAATTCAGCAGTTGTTAGAAATGGTCTCTTTTATCAACAGTAACCTGAAATTTGCA 2488
QY 740 GluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPheLysLeuPhe 759
DB 2489 GAGGCACCATCATGCTGTGATTTTTCAGATGCTCGATTGGAAAAGACTTTTAAACTATTT 2548
QY 760 LysLysIlePheProSerLeuGluLeuAasnIleThrAspLeuLeuGluAtpThrProArg 779
DB 2549 AAAAAATTTTCTCTCTGGAATTAATATAACAGATTTACTTTGAAGACACTCCCAGA 2608
QY 780 GlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyrAspAspPro 799
DB 2609 CAGTCCGGATATGTGGAGGCTTGCAATGTATGAGTGTAGAGAATGCTACGACGATCCG 2668
QY 800 AspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAasnThrGlnValHisLeu 819
DB 2669 GACATCTCAGCTGGAAAAATCAAGCAGTTTTGTAAAAACCTGCAACACTCAAGTCCACCTT 2728
QY 820 HisProLysArgLeuAasnHisLysTyrAsnProValSerLeuProLysAspLeuProAsp 839
DB 2729 CATCCGAAGAGGCTGAATCATATAAATAAACCCAGTGTCACTTCCCAAGACTTACCCGAC 2788
QY 840 TrpAspTyrArgHisGlyCysIleProCysGlnAasnMetGluLeuPheAlaValLeuCys 859
DB 2789 TGGGACTGGAGACACGCGTCATCCCTTGGCAGAAATATGAGATTTATTTGCTGTTCTCTGC 2848
QY 860 IleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSerAlaTyrLeu 879
DB 2849 ATAGAAACAGCCACTATGTTGCTTTTGTGAAGTATGGGAAGGACGATTTCTGCTGGCTC 2908
QY 880 PheAspSerMetAlaAspArgAspGlyGlyGlnAasnGlyPheAenIleProGlnVal 899
DB 2909 TTTCTTGACAGCATGCGGATCGGATGGGTGAGTTCAGAAATGGCTTCAACATTTCTCAAGTC 2968
QY 900 ThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAtpLeuHisSerLeu 919

```

Db      2969  ACCCATGCCAGAGTAGGAGACTTGAAGATGCTCTCGAAGACCTGCAATTCCTTG 3028
Qy      920  AspSerArgArgileGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyrMetCysMet 939
Db      3029  GACTCCAGAGAAATCCAAAGCTGTGCACGAAGACTGCTTTGTGATGCATATATGTGCATG 3088
Qy      940  TyrGlnSerProThrMetSerLeuTyrLys 949
Db      3089  TACCAGAGTCCCAACATGAGTTGTGTACAAA 3118

RESULT 9
LOCUS   AKI22389          4314 bp  mRNA  linear  ROD 15-MAR-2003
DEFINITION Mus musculus mRNA for mKIAA0849 protein.
ACCESSION AKI22389
VERSION   AKI22389.1  GI:28972434
KEYWORDS  FLI_CDNA.
SOURCE    Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE
AUTHORS   Okazaki,N., Kikuno,R., Ohara,R., Inamoto,S., Aizawa,H., Yuasa,S.,
          Nakajima,D., Nagase,T., Ohara,O. and Koga,H.
TITLE     Prediction of the coding sequences of mouse homologues of KIAA
          gene: II. The complete nucleotide sequences of 400 mouse
          KIAA-homologous cDNAs identified by screening of terminal sequences
          of cDNA clones randomly sampled from size-fractionated libraries
          DNA Res. 10, 35-48 (2003)
JOURNAL  2 (bases 1 to 4314)
REFERENCE Okazaki,N., Kikuno,R., Nagase,T., Ohara,O. and Koga,H.
AUTHORS   Direct Submission
TITLE     Submitted (07-FEB-2003) Hiesashi Koga, Kazusa DNA Research
          Institute, Laboratory for Genome Informatics, 2-6-7
          Kazusa-kamatari, Kisarazu, Chiba 292-0818, Japan
          (E-mail:mousekazusa.or.jp, Tel:81-438-52-3919, Fax:81-438-52-3918)
          The CRATE program supported by Japan science and technology
          corporation; cDNA full insert sequencing: Kazusa DNA Research
          Institute; cDNA library construction, clone selection and 5'- &
          3'-end one pass sequencing.
FEATURES             Location/Qualifiers
     1..4314
     /organism="Mus musculus"
     /mol_type="mRNA"
     /db_xref="taxon:10090"
     /clone="mbh04492"
     /tissue_type="brain"
     /dev_stage="adult"
     /notes="vector:modified pBC SK+"
     1..4314
     /gene="mKIAA0849"
     /evidence=not_experimental
     /product="mKIAA0849 protein"
     /protein_id="BAC65671.1"
     /db_xref="GI:28972435"
     /translation="IAVELQHGCOVAKILLMDLQVRHFYDRFTSGSFLPVNVLKV
     TVTMSGLSQEAKVTSFYHEERIFXLLQECSTVDKOTOKLLKVPKSGISGOYIQRDSV
     GHSRVSPTGKNQIGKILEQHPAVLFVDKVDVINEKFTPELLLAIINCEERLSLF
     RNRRLSKGLQVGVSKVQLRSBEEKFPVGRPLLAERTVSGIFPGVLLSEB
     RGQFTDGVYQKQLFCQCDGFVALDKLEIEDDNDGLESDFDAGPDMQVBPFP
     LEINRSVLKVGSESTGTVIFCDVLPKGESLGYFVGVDMNDPIGNWDRGDFQVQLCS
     FASVESTILLHINDIIPDSVTQERRPKLAFMSRGVGDKGSSSHNPKPKVTGSDPSG
     RNRSELFTYLINGSVDSQSKSKPNVYIDEVADPAKSLTMSNSDCHSPPPQPSM
     NLSSENTHSLSPFLTKPNTINGSMAHSPLSLVSQSVGMGLNSTPVPQSPPLPISG
     NAHGLVGSLSAEYKENPPFVGIIRWIGQPPGLSDVLAGELEDECACTDGTGRTGTY
     FTCAKALKALFKCRPSRFASLPQVSNQIERCNLSLAPFGYLSVEVEENTPPKMEK
     EGLEINIGKKGKIQGHVNSCYLDSLFLCLFAPFSALDVLRLPKKNDIEYSETQEL
     LRTEIVNPLRIYGVVCAATKIMLKILEKVEAASGFTSBEKDPPEEFNLINILPHDLIRVE

```

```

PLKIRSAGQKQVDCNFYQIFMEKNEKVGVPITIQQLLENSFINSNLKFAEAPSCLLIQ
MPRFKDFKFKKIFPSELNITDLLEDTPROCRICGGLAMYEKRECYDPAVLISAGKI
KQFCCTKSTQVHLHPRRLNHSYHPSLKPDLMDWRHGCICPQKRMFLAVLCIETSH
YVAFVKYGGKDSAWLFFDSMADRDGCGFNIPQVTPCPVEGVSYLKMSLDLHLSDSR
RTQGCARRLLCDAVWCMYQSPMTSLYK"

ORIGIN
Alignment Scores:
Pred. No.:      0      Length:      4314
Score:          4718.50      Matches:      898
Percent Similarity: 96.54%      Conservative: 22
Best Local Similarity: 94.23%      Mismatches: 28
Query Match:      93.73%      Indels: 5
DB:              10      Gaps: 5

US-09-671-687A-3 (1-949) x AKI22389 (1-4314)

Qy      1  MetSerSerGlyLeuTyrPsrSerGlnGluLysValThrSerProTyrTrpGluGluArgile 20
Db      338  ATGAGTTTCAGGCTGTGGAGCCCAAGAGAAGTTACTTCACTCCCTACTGGGAAGACCGGATT 397
Qy      21  PheTyrLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLeuLys 40
Db      398  TTTTATCTCTCTTCTTCAAGATGCAGTGTAAACACAACTCAAGAGCTGCTGAAA 457
Qy      41  ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgile 60
Db      458  GTACCCAAAGGAGTATAGGACAGTACATCCAAAGACCGTTCTGTGGGCATTCAGAGTT 517
Qy      61  ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGlnGlnProHisAla 80
Db      518  CCTTCCACAAAAGGCAAGAAAATCAGATTGGATTAAAAATCTTGGAGCAACCCGATGCA 577
Qy      81  ValLeuPheValAspGlu---AspValValGluLeuAsnGluLysPheThrGluLeuLeu 99
Db      578  GTTCTCTTTTGTGATGAAAAGGATGTTGTAGAAATAAATAAATAATTCACAGAGTTACTG 637
Qy      100  LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
Db      638  TTGGCAATTACCAACTGTGAGGAGAGGCTCAGGCTATTATAGAACACAGACTCCGACTAAGT 697
Qy      120  LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
Db      698  AAAGGCTCCAGGTAGACGTCGGCAGTCTCTGTGAAAGTACAGCTCGCATCTGGGGAAGAG 757
Qy      140  LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
Db      758  AAATTTCCAGGAGTTGTACGCTTCAGAGGACCTTTATTAGCGAGAGACCGTGTCCGGG 817
Qy      160  IlePhePheGlyValGluLeuLeuGluGlyArgGlyGlnGlyPheThrAspGlyVal 179
Db      818  ATTTTCTTTGGAGTAGAATTTTGGAGAAGGTCGTGCTCAAGTTTCACGATGGGGTA 877
Qy      180  TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
Db      878  TATCAAGGGAAGCAGACTTTTCCAGTGTGATGAGGACTGTGGCGGTGTTTGTTCATTGGAC 937
Qy      199  LysLeuGluLeuLeuGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
Db      938  AAATCTAGAACCTTAGAAGATGATGACAAATGGATTGGAAAGTGAATTTTGCAGGCCCAAGGA 997
Qy      219  AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly 238
Db      998  GATACAAATGAGGTTGAACTCCCTCCCTTTGGAAATAAATCCACAGAGTTCTTTGAGAGTT 1057
Qy      239  GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
Db      1058  GGAGAAAGTACAGAACTCTGGAACAGTAATATCTGTGATGTTTACCAGGAAAAAGAGAGT 1117
Qy      259  LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTyrAspGlyArgPhe 278
Db      1118  CTAGGATATTGTTGTTGGTGGACATGAGTAACCTATTGGCAACTGGGATGGAGAGGTTT 1177

```

QY 279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297.
DB 1178 GATGGAGTACAGCTCTGTAGTTTGGCAAGTGTGAAGAGTACAATCTCTCGACATCAAT 1237
QY 298 AspIleIleProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMetSer 317
DB 1238 GACATCATCCAGATAGCTGCACAGAGAAAGAGGAGGCTCCCAAATCTTGCCTTTATGTCA 1297
QY 318 ArgGlyValGlyAspLysGlySerSerSerHisAsnLysProLysAlaThrGlySerThr 337
DB 1298 AGAGGTGTAGGTGACAAAGGTTTCATCTAGTACATAAATACCAAGGTTACAGGATCTACC 1357
QY 338 SerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVal 356
DB 1358 TCAGACCCCTGGAAGTGAAGACAGATCTGAATTAATTTATACCTTAATGGTCTATCTGT 1417
QY 357 AspSerGlnProGlnSerLysSerLysAsnThrTyrTyrIleAspGluValAlaGluAsp 376
DB 1418 GACTCACAA---CAATCCAGTCCCAAAATCCATGGTACATTGATGAGGTTGCAGAGAC 1474
QY 377 ProAlaLysSerLeuThrGluLeuSerThrAspPheAspArgSerSerProProLeuGln 396
DB 1475 CTTGCAAAAGTCACTTACAGAGATGCTCTCGGACTTCGGACATTCATCTCTCTCCACCGCAG 1534
QY 397 ProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeu 416
DB 1535 CCTCTTCCATGAATCTCTGTCTAGCAGAGACAGATTCACCTCTTACCTCTTACGCTG 1594
QY 417 ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuAlaGln 436
DB 1595 ACAAGATGCCCACTACTAATGGCAGCATGGCTCATAGTCCACTCTCTGTCTGTCAGTGAC 1654
QY 437 SerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPro 456
DB 1655 TCTGTGATGGGGAGCTGACAGACACACTGTCTCAGAGAGTCCACCTTGGCCATCTCT 1714
QY 457 ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro 476
DB 1715 TCTGGGAATGCACACGGCTAGAGGTGGCTCACTGGCTGAGTAAAGAGAACCCCGG 1774
QY 477 PheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGly 496
DB 1775 TTCTATGGGGTTATCCGTTGGATGTCGCCAGCCACCGAGGCTCAGTACGCTGCTAGTGA 1834
QY 497 LeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr 516
DB 1835 CTGGAACTGGAGATGAATCGCAGGCTGTACAGATGGAACCTTCAGGGGCGCCGGTAT 1894
QY 517 PheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSer 536
DB 1895 TTCACGTGTGCCCTGAAGAAGGCACCTGTTGTGAACCTGAAGAGCTGCAGACCGGACTCT 1954
QY 537 ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe 556
DB 1955 AGGTTTGCATCTTGCAGCCTGTTTCCAATCAGATTGAAAGGTGTAACTCTTTAGCATTT 2014
QY 557 GlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluGly 576
DB 2015 GGGGGCTATTTAAGTAGAGTAGTAGAGAAATTAACCTCAATTAATGGAAGGAGGAGGT 2074
QY 577 LeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu 596
DB 2075 TTAGAGATTAATGTTGAAGAAGAAAGGATCCAGGGCCATTACAACTTCTTGTACTTA 2134
QY 597 AspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArg 616
DB 2135 GACTCACTTATCTGCTATTGCTTTTGTCTTGTCTGCTGACACTGTGTTACTTGA 2194
QY 617 ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu 636
DB 2195 CCCAAAGAGAAATGATATAGAGTATTACAGTGAGACTCAGGAGCTACTGAGACAGAG 2254
QY 637 IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg 656

DB 2255 ATAGTCAATCCTCTGAGATATATGATATGTGTGTCACCAAGATTTGAAACTGAGG 2314
QY 657 LysIleLeuGluLysValGluAlaSerGlyPheThrSerGluGluLysAspProGlu 676
DB 2315 AAAATACTTGAAGAGTGGAGGCTGCATCAGGTTTACCTCTGAAGAAAAAGATCTCGAA 2374
QY 677 GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArg 696
DB 2375 GAATTTCTAAATATCTCTGTTTCATGATATTTAAAGGGTTGAACCATTTGTTAAAAATAAGA 2434
QY 697 SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu 716
DB 2435 TCAGCAGGTCAAAAGTTCAGACTGTAACTCTTATCAAAATTTTATGGAAGAAAAATGAG 2494
QY 717 LysValGlyValProThrIleGlnGlnLeuLeuGluTrpSerPheIleAsnSerAsnLeu 736
DB 2495 AAAGTTGGAGTACCCACAATTCAGCAGTTATTAGATGGTCTTTTATCAACAGCAACCTG 2554
QY 737 LysPheAlaGluAlaProSerCysLeuIleGlnMetProArgPheGlyLysAspPhe 756
DB 2555 AAATTTGCAGAGCACCATCATCTGATTATCCAGATGCTCGGTTTGGAAAAAGACTTT 2614
QY 757 LysLeuPheLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAsp 776
DB 2615 AAATCTATTAAAAAAATTTTCTCTCCCTGGAATTAATATAACAGATTTACTTTGAAGAC 2674
QY 777 ThrProArgGlnCysArgIleCysGlyLeuAlaMetTyrGluCysArgGluCysTyr 796
DB 2675 ACTCCAGCAGCTGCGCATCTGTGGAGGACTTGCAGTGTACGAGTGCAGGGAGTGTAT 2734
QY 797 AspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGln 816
DB 2735 GACGATCCGAGACATCTCAGCTGGGAAGATCAAGCAGTTCTGTAAAGACTGCGACACTCAG 2794
QY 817 ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAsp 836
DB 2795 GTTCACCTTCATCCAGGAGGTTGAATCATCTTATCATCCAGTATCATCTCCCAAGAC 2854
QY 837 LeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAla 856
DB 2855 TTCCCTGACTGGGACTGGAGACAGGCTGCATCCCTGTGAGAGATGGAGTATTTGCT 2914
QY 857 ValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspSer 876
DB 2915 GTTCTCTGCATAGAACGACCACTATGTTGCTTTTGTGAAGTATGGAGGATGACTCT 2974
QY 877 AlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIle 896
DB 2975 GCCTGGCTTTCTTTGACAGCATGGGGATCGAGATGGTGGTTCAGAAATGGCTTCAACATT 3034
QY 897 ProGlnValThrProCysProGluValGlyGlyTyrLeuLysMetSerLeuGluAspLeu 916
DB 3035 CCACAAAGTGAAGCTTCCAGAGTGGGAGAGTACTTGAAGATGTCTCTGGAGGACCTG 3094
QY 917 HisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyr 936
DB 3095 CACTCTTGGACTCCAGAGGATTCAGGCTGTGGCGCAGACTTCTTTGGATGCATAC 3154
QY 937 MetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
DB 3155 ATGTGCATGTACCAGAGTCCACCATGAGCTGTACAAA 3193
RESULT 10
BC042438
LOCUS
DEFINITION
Mus musculus cylindromatosis (turban tumor syndrome), mRNA (cdna
clone MGC:25429 IMAGE:3983771), complete cds.
ACCESSION
BC042438
VERSION
BC042438.1
KEYWORDS
MGC.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus

Db 776 AACTAGAACTTATAGACGATGATCAATGGATTGGAAGTGATTTTGCAGGCCCCAGGA 835
Qy 219 AspThrMetGlnValGluLeuProLeuGluLeuAAsnSerArgValSerLeuLysGly 238
Db 836 GATACAAATGACAGTTGAACCTCCCTCTTGGAAATAAATCCACAGAGTTCTTTGAAGGTT 895
Qy 239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
Db 896 GGAGAAAGTACAGAAATCTGGAACAGTAATATCTGTGATGTTTACAGGAAAAGAGAT 955
Qy 259 LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe 278
Db 956 CTAGGATATTTGTTGTTGGTGGACATGGATAACCTATTGGCAACTGGGATGGAAGGTTT 1015
Qy 279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
Db 1016 GATGGAGTACAGCTCTGTAGTTTGGCAAGTGTGAAAGTACAATCTCTCGCACATCAAT 1075
Qy 298 AsnIleIleProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMetSer 317
Db 1076 GACATCATCCACAGATAGCTGACAGGAAAGAGGCGCTCCCAAACTTGCCCTTTATGTCA 1135
Qy 318 ArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThrGlySerThr 337
Db 1136 AGAGGTGAGTGCACAAAGTTTCATCTAGTCATAATAACCAAGGTTACAGGATCTACC 1195
Qy 338 SerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVal 356
Db 1196 TCAGACCTCGGAAGTAGAACAGATCTGAATATTTTATACCTTAATAATGGTTCATCTGTT 1255
Qy 357 AspSerGlnProGlnSerLysSerLysAsnThrTyrIleAspGluValAlaGluAsp 376
Db 1256 GACTCACAA---CAATCCCAAGTCCAAAATCCATGGTACATTTGATGAAGTTGCAGAGAC 1312
Qy 377 ProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerProProLeuGln 396
Db 1313 CTGCAAGTCACTTACAGAGATGTTCTCGGACTTCGACATTCATCTCTCCACCGCAG 1372
Qy 397 ProProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeu 416
Db 1373 CCTCTTCCATGNACTCCTTGCTAGCGAGAACAGATTCACCTCCTTACCTTCAGCCTG 1432
Qy 417 ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuAlaGln 436
Db 1433 ACAAGATGCCAAATCAATGACGATGCTCATAGTCCACTCTCTCTGTGTCAGTGCAG 1492
Qy 437 SerValMetGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPro 456
Db 1493 TCTGTATGGGGAGCTGAACAGCACACCTGTCCAGGAGTCCACCTTGCCCATCTCT 1552
Qy 457 ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro 476
Db 1553 TCTGGGAATGCACACGGCTAGAGTGGGCTCACTGGCTGAATGAAGAAACCCCGG 1612
Qy 477 PheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGly 496
Db 1613 TTCTATGGGTTATCCGTTGGATTGGCCAGCCACCGGCTCAGTGACGTGCTAGTCTGA 1672
Qy 497 LeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr 516
Db 1673 CTGGAATCTGGAAGATGAATGCGCAGGCTGTACAGATGGAACTTTTCAGGGGCGCGGTAT 1732
Qy 517 PheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSer 536
Db 1733 TTCACGTGTGCCCTGAAGAGGACACTGTTTGTGAACCTGAAGAGCTGCAGACCGGACTCT 1792
Qy 537 ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe 556
Db 1793 AGGTTTGCATCCTTGACGCTGTTTCCATCAATGAAGGTGTAATCTCTTTAGCATTT 1852
Qy 557 GlyGlyTyrLeuSerGluValValGluAsnThrProProLysMetGluLysGluGly 576
Db 1853 CGGGGCTATTTAAGTGAAGTAGTAGAAGAAATACTCCACCTAAATGGAAAGGAAGGT 1912

Qy 577 LeuGluIleMetIleGlyLysLysLysGlyIleGlnGlyHisTyrAsnSerCystTyrIleu 596
Db 1913 TTAGAGATAATGATTTGGAAAGAAAGAGGATCCAGGGCCATTACAATCTTGTGTACTTAA 1972
Qy 597 AspSerThrLeuLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArg 616
Db 1973 GACTCAACTTATCTGCTTATTTGCTTTTAGTTCTGCCCTGGACACTGTGTACTTAA 2032
Qy 617 ProLysLeuLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu 636
Db 2033 CCCAAAGAGAAGAAATGATATAGAGTATACAGTCAGACTCAGGAGCTACTCAGGACAGAG 2092
Qy 637 IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg 656
Db 2093 ATAGTCAATCTCTGAGATATATGGATATGTGTGTGCCCAAAAAGATTATGAACATGAGG 2152
Qy 657 LysIleLeuGluLysValGluAlaAspGlyPheThrSerGluGluLysAspProGlu 676
Db 2153 AAAATCTTGAAGAGTTGAGGCTGCATCAGTTTTTACCTCTGAAGAAAAGATCCTGAA 2212
Qy 677 GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLysIleArg 696
Db 2213 GAATCTTAAATATCTGTTTTCATGATATTTAAAGGTTGAACCATCTGTTAAAATAAGA 2272
Qy 697 SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu 716
Db 2273 TCAGCAGGTCAAAAGTTCAAGACTGTAACTTCTATCAAAATTTTATGGAAAAAATGAG 2332
Qy 717 LysValGlyValProThrIleGlnGluLeuGluTrpSerPheIleAsnSerAsnLeu 736
Db 2333 AAAGTTGGAGTACCCAAATTCAGCAGTATTAAGATGGTCTTTTATCAACAGCAACCTG 2392
Qy 737 LysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPhe 756
Db 2393 AAATTTGCAGAGGCACCATCATGCTTGATTTATCCAGATGCCTCGGTTTGGAAAGACTTT 2452
Qy 757 LysLeuPheLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAsp 776
Db 2453 AAACATTTAAAAAATTTTCTCTCCCTGGAATTAATAATAACAGATTTTACTTTGAAGAC 2512
Qy 777 ThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyr 796
Db 2513 ACTCCAGGAGTCCGCATCTGTGGAGACTTCGGATGTACGAGTGCAGGAGTGCTAT 2572
Qy 797 AspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGln 816
Db 2573 GACATCCGACATCTCAGCTGGGAAGATCAAGCAGTTCCTGTAAAGCTTGAGCACTCAG 2632
Qy 817 ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAsp 836
Db 2633 GTTCACTTCATCCAGAGGTTGAATCATTTTATCATCCAGTATCATCTTCCCAAGAC 2692
Qy 837 LeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAla 856
Db 2693 TTGCCTGACTGGGACTGGAGACGCGTGCATCCCTGTCTGCAAGAGATGGAGTTTCTCT 2752
Qy 857 ValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSer 876
Db 2753 GTTCTCTGCATAGAAACGACCATATGTGCTTTTGTGAAGTATGGGAAGATGACTCT 2812
Qy 877 AlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIle 896
Db 2813 GCCTGGCTTTTCTTTGACAGATGGCGGATCGAGATGGTGGTCAAGATGGCTTCAACAT 2872
Qy 897 ProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLeu 916
Db 2873 CCACAGTGCAGCCTGCCCAAGAGTGGAGAGTACTTGAAGATGCTCTCTCGAGGACCTG 2932
Qy 917 HisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyr 936
Db 2933 CACTCTTTGGATTCCAGAAAGATTCAAGGCTGTGCGCGCAGACTTCTTTTGGATGCATAC 2992

```

QY 937 MetCysMetTyGlnSerProThrMetSerLeuTyrLys 949
Db 2993 ATGTGATGTACAGAGTCCAAACCATGAGCGCTGTACAAA 3031

RESULT 11
BC082001 3267 bp mRNA linear ROD 02-SEP-2004
LOCUS Rattus norvegicus cdna clone MGC:94236 IMAGE:7131198, complete cds.
DEFINITION BC082001
ACCESSION BC082001
VERSION BC082001.1 GI:51858716
KEYWORDS MGC.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus
1 (bases 1 to 3267)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Schetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S.S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahy,J., Helton,E., Kettman,M., Madan,A., Rodrigues,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Boutfard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalusz,D.E.,
Schnerch,A., Schein,J.E., Jones,S.U. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 3267)
Director MGC Project.
Direct Submission
Submitted (01-SEP-2004) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapsb@mail.nih.gov
Tissue Procurement: Howard Jacobs
CDNA Library Preparation: Express Genomics
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate; 194 Row; J Column; 24
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, Similarity but not identity to protein.
Location/Qualifiers
1. .3267
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="MGC:94236 IMAGE:7131198"
/tissue_type="Testis, rat (Brown Norway)"
/clone_lib="NIH_MGC_238"

FEATURES
source

```

```

/lab_host="DH10B"
/notes="Vector: pExpress1"
196. .3057
/codon_start=1
/protein_id="AAH82001.1"
/db_xref="GI:51858717"
/translation="MSSGLWNEKVTSPYWEERLFYLLLOECSSVTDKQTKLRVPKGS
STGQYIDRSVGHRSKAGKQKQIGLKEQHAFLVDEKDVVEINKEFTTELLIA
ITNCEERLSLFRNRLSKGLVDVQSVLRSGEEKFPVGVRRPGPLLABRTVSG
IEFVVELLEGRCQGTGQYQKQLFQCDGCGVFVLDKLELIEDDDNGLSDFAE
PGDTVQVEPPPLEINSRSLKVGSESTGVTFCDVLPKGSGLGTFVGVDDNPNFKNM
DGRFQVQLCSFASVESTVLLHNDIIPDSVTQERRPPKLAFLMRGVDGSSHNPK
KVTGSTSDPSRNRSSELYFLANGSSVDQSKNPNWYDEVAEDPAKSLTMSDDP
GHSSPPQPPNNLSLSENRHSLPFLSKPMPTNGSMASPLSLSVSGVGLNLTSP
VOESPMPSSSGNAHGLEVAENKPPFYGVIRWIGOPGGLSDVLAGLEDECA
GCTDGTFRGTRYTCALKALFKVLSKCRDPSRPSLOPVNSQIERCNSLAFGLYSE
VVEENTPPKWEKEGLIMIKKGIQHNCSYDSDTLFCIFAFSSALDVLVLRPKR
NDVEYSETOELLRTIEIVNPLRIYGVYCATIKMKRILKLEKVAASGDTSEEDKPEF
LNILFHDLRVLPKLRAGQVQDCNFYQIFEMKNEKVGVTIQQLLEWSTFNSL
KFAEAPDCLIIOMPRGKDFPKFKIPPSLELNITDLEDTPRQCRICGGLAMYECRE
CYDDPDISAGKIQCFKCTCTOVHLHPRLNHTYHPVSLPKDLPMMDWRHGCI PCQKM
ELFVLCIETSHYVAFVKYKDDSAWLFPDSMADRGQGNFPIQVTPCPVEGYLK
MSLEUHSUDSRRIQGCARRLLCDATMCMYQSTMSLYK"
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3267
Score: 4712.00 Matches: 893
Percent Similarity: 96.54% Conservative: 27
Best Local Similarity: 93.70% Mismatches: 29
Query Match: 93.60% Indels: 4
DB: 10 Gaps: 4
US-09-671-687A-3 (1-949) x BC082001 (1-3267)
QY 1 MetSerSerGlyLeuTyrSerGlnGluLysValThrSerProTyrTyrGluGluArgile 20
Db 196 ATGAGTTTCAGGCTGTGGAACCAAGAGAAAGTTTACTTCCCTACTCGGAAGACGGCTT 255
QY 21 PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLys 40
Db 256 TTTTATCTGCTTCTTCAAGAAATGCAAGTGAACAGACAAACAGACACAGAAAGCTCTCTGAGA 315
QY 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgile 60
Db 316 GTACCCAAAGGAGCATAGACAGTACATCCAGACCGTTCCTCGGGGCATTCAGAGTT 375
QY 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla 80
Db 376 CCTTCTGCTAAAGCAAGCAAAAAATCAGATTGGATTAAAAATCTTAGAGCAACCGCATGCA 435
QY 81 ValLeuPheValAspGlu---AspValValGluIleAsnGlnLysPheThrGluLeuLeu 99
Db 436 GTTCTGTGTTGTTGATCAAAAGGATGTTGTGAAAAATAAATGAAAAATTCACAGAGTTACTG 495
QY 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
Db 496 TTGGCAATTACCAACTGTGAGGAGAGGCTCAGCCCTATTATTAGAAACAGAAATCCGACTAAGT 555
QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
Db 556 AAAGGCTCCAGGTAGACGTGGGAGTCTCTGTGAGAGTACAGCTCGCATCTGGGGAGGAG 615
QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
Db 616 AAGTTTCCAGGAGTTGTACGCTTCAGAGGACCTTTATTAGCGGAGAGACGGTGTGCGGG 675
QY 160 IlePhePheGlyValGluLeuLeuGluGluGlyArgGlyGlnGlyPheThrAspGlyVal 179
Db 676 ATTTTCTTTGGAGTAGAATTTACTGGAAGAGAGTCTGCGCCAGGTTTTCACATGATGGGTG 735
QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198

```


Db	736		TATCAGGAAACAGCTCTTCCAGTGTGATCAGACCTGGCGGTTTTTGTTGCAATGGAC	795
Qy	199	LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly		576
Db	796	AAGCTGGAGCTATAGAACATGATGACAAATGGGTGGAAAGTGATTTTGCAGGCCCAGGA	855	1876
Qy	219	AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly		596
Db	856	GATACAGTCCAGGTGAACCTCCCTTTGGAAATAAATCCAGAGTTTCTTTGAAGGTT	915	1936
Qy	239	GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer		616
Db	916	GGAGAAAGTACAGAACTGGACAGTGATATCTCTGATGTTTTACAGGAAAAAGAGT	975	1995
Qy	259	LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe		636
Db	976	CTAGGATAATTTGTTGGTGGACATGGATAACCTATTGGCAACTGGGACGGAAGGTTT	1035	2055
Qy	279	AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn		656
Db	1036	GATGGGTTTCAGCTTTGCAAGTTTGGCAAGTGTGGAGTACAGTCTCTCTACACATCAAT	1095	2116
Qy	298	AspIleIleProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMetSer		676
Db	1096	GACATCATCCAGATAGCGTGACACAGGAAGAGAGACCTCCCAAACTTGCCTTTATGTCA	1155	2175
Qy	318	ArgGlyValGlyAspLysGlySerSerSerHisAsnLysProLysAlaThrGlySerThr		696
Db	1156	ACAGGTGTAGGTGACAAAGTTTCATCTAGTCATAATAACCAAGGTTACAGGATCTACC	1215	2235
Qy	338	SerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVal		716
Db	1216	TCAGACCTCGGAAGTAGAACAGATCTGAATATATTTATACCTTAAATGGGTCTCTGTT	1275	2295
Qy	357	AspSerGlnProGlnSerLysSerLysAsnThrTyrIleAspGluValAlaGluAsp		736
Db	1276	GACTCACAAACAATCCAAAGTCCAAAAACCCATGGTACATTGATGAAGTTGCAGAGAC	1335	2355
Qy	377	ProAlaLysSerLeuThrGluIleSerThrAppPheAspArgSerProProLeuGln		756
Db	1336	CTGCAAGTCACTTACAGAGATGCTTCAGACTTCGGACATTCAGCTCCACCGCAA	1395	2415
Qy	397	ProProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeu		776
Db	1396	CCTCCTTCATGAATCTCTGTAGCGAGAACAGATTCACACTCTTACCTTCAGCCTG	1455	2475
Qy	417	ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuAlaGln		796
Db	1456	ACAAAGATGCCCAACACTAATGGCAGCATGGCTCACAGTCCACTCTCTCTGTCAGTGCAG	1515	2535
Qy	437	SerValMetGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPro		816
Db	1516	TCTGTGATGGGAGAGCTGAACAGACTCTCTGTCCAGAGAGTCCACCCATGCCAGCTCT	1575	2595
Qy	457	ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro		836
Db	1576	TCCTGGAAATGCACCGGCTAGAGTGGGCTCACTGGCTGAAGTAAAGAGAACCCCGG	1635	2655
Qy	477	PheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGly		856
Db	1636	TTCTATGGGTTATCGTTGGATTGGCCAGCACCCAGGCTCAGTACGCTGCTTGTCTGA	1695	2715
Qy	497	LeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr		876
Db	1696	TTGGAACCTGGAAAGATGAATCGCAGAGTTGTACGGATGGAACTTTACGGGGCAGCGCTAT	1755	2775
Qy	517	PheThrCysAlaLeuLysIleValLeuPheValLysLeuLysSerCysArgProAspSer		896
Db	1756	TTCACTGTGGCTTGAAGAAAGACACTGTTCTGGAAACTGAAGAGCTGCAGACCACTCT	1815	2835
Qy	537	ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe		916
Db			2955	2895
Db	1816	AGGTTTGCATCCTTGCAGCCTGTTTCCAAATCAGATCGAAAGGTGTAACCTTTTACCATTT	1875	
Qy	557	GlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluGly	576	
Db	1876	GGGGGCTTACTTAAGTGAAGTAGTAGAAGAAAATACGCCACCTAAATTCGAAAAGGAGGT	1935	
Qy	577	LeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu	596	
Db	1936	TTAGAGATAATGATTTGGAAAGAAAGAGGATCCAGGGCCATTACAAATCTTGTGTACTTAA	1995	
Qy	597	AspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArg	616	
Db	1996	GACTCAACTTTATCTGCTTATTTGCTTTTAGTTCTGCCCTGGACACTGTATTTACTTAA	2055	
Qy	617	ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu	636	
Db	2056	CCCAAGAGAGAAGATGATGTAGAGTATTACAGTGAGACTCAAGAGCTACTGAGGACAGAG	2115	
Qy	637	IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg	656	
Db	2116	ATAGTCAATCCTCTGAGAATATATGGATATGTGTGTGCCCAAAAGATTATGAAGCTGAGG	2175	
Qy	657	LysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGlu	676	
Db	2176	AAATACTTGAAGAAAGTTGAGGCTGCATCAGGATTTACCTCTGAGGAAAAAGATCCTGAA	2235	
Qy	677	GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArg	696	
Db	2236	GAATTTCTAAACATCCTGTTTCATGATATTTTAAGGGTTGAACCATTTGTTAAAAATAAGG	2295	
Qy	697	SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu	716	
Db	2296	TCAGCAGGTCAAAAGTTCAAGACTGTAACTTCTATCAAAATTTTATGGAAAAAATGAG	2355	
Qy	717	LysValGlyValProThrIleGlnLeuLeuGluTrpSerPheIleAsnSerAsnLeu	736	
Db	2356	AAAGTCGGAGTACCCCAATCCAGCAGTTATTAAGATGGTCTTTTATCAACAGCAACCTG	2415	
Qy	737	LysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPhe	756	
Db	2416	AAATTTGCAGAGGACCATCATGCTTGATTTATCCAGATGCTTCGGTTTGGGAAGACTTT	2475	
Qy	757	LysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAsp	776	
Db	2476	AAACTATTTAAAAAATTTTCTCTCCCTGGAATTAATAATAACAGATTTACTTTGAAGAC	2535	
Qy	777	ThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyr	796	
Db	2536	ACTCCAGGCGAGTCCCGCATCTGTGGAGGACTCGCCATGTATGATGTGTAGAGGTGCTAT	2595	
Qy	797	AspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGln	816	
Db	2596	GATGACCCGACATCTCGGACGGAGAGATCAGCAGTTCTGTAAAGCTTGAGACTGAGCTCAG	2655	
Qy	817	ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAsp	836	
Db	2656	GTTCACTTCATCCAGAGACTGAATCACACTTACCATCCAGTATCACTTCCCAAGAC	2715	
Qy	837	LeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAla	856	
Db	2716	TTGCCGAGCTGGGACTGGAGACACGGCTGCATCCCGTGTGAGAAAGATGGAGTTTATGCT	2775	
Qy	857	ValLeuCystIleGlnThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSer	876	
Db	2776	GTCTCTGCATAGAAACCAAGCCACTATGTGTCTTTGTGAAGTACGGGAAGATGAGCTCT	2835	
Qy	877	AlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIle	896	
Db	2836	GCCTGGCTCTCTTTGACAGCATGGCTGATCGAGATGCTGGTCAAGATGGCTTCAACATT	2895	
Qy	897	ProGlnValThrProCysProGluValGlyGlyLeuLysMetSerLeuGluAspLeu	916	
Db	2896	CCCAAGTGACACCTCGCCCAAGAGTAGGAGAGTACTTGAAGATGTCTCTGAGGAGACCTG	2955	

QY 917 HisSerLeuAspSerArgArgIleGlnGlyCysAlaArgLeuLeuCysAspAlaTyr 936
 Db 2956 CACTCTTTGGACTCCAGAGAGATTCAAGGCTGTGCGCGAGACTTCTTTGGATGCATAC 3015

QY 937 MetCysMetTyrClnSerProThrMetSerLeuTyrIys 949
 Db 3016 ATGTGCATGTACAGAGTCAACCATGAGCTGTGTACAAA 3054

RESULT 12
 BD160617 2845 bp DNA linear PAT 17-JAN-2003
 Locus
 DEFINITION Primer for synthesizing full-length cDNA and use thereof.
 ACCESSION BD160617
 VERSION BD160617.1 GI:27866375
 KEYWORDS JP 2002191363-A/15460.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 2845)
 Ota.T., Isogai.T., Nishikawa.T., Hayashi.K., Saito.K., Yamamoto.J.,
 Ishii.S., Sugiyama.T., Wakamatsu.A., Nagai.K. and Otsuki.T.
 Primer for synthesizing full-length cDNA and use thereof
 Patent: JP 2002191363-A 15460 09-JUL-2002;
 JOURNAL HELIX RESEARCH INSTITUTE
 COMMENT OS Homo sapiens (human)
 PN JP 2002191363-A/15460
 PD 09-JUL-2002
 PF 28-JUL-2000 JP 2000280990
 PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
 PI SAITO,
 PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
 PI KEIICHI NAGAI, TETSUJI OTSUKI
 PC
 C12N15/09, C07K14/47, C07K16/18, C12N1/19, C12N1/21, C12N5/ PC
 10,
 PC C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
 Primer for synthesizing full-length cDNA and use thereof PH Key
 Location/Qualifiers
 FT CDS Location/Qualifiers
 1..2845
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

ORIGIN
 Alignment Scores:
 Pred. No.: 0 Length: 2845
 Score: 4601.00 Matches: 875
 Percent Similarity: 99.32% Conservative: 0
 Best Local Similarity: 99.32% Mismatches: 2
 Query Match: 91.40% Indels: 4
 DB: 6 Gaps: 4

US-09-671-687A-3 (1-949) x BD160617 (1-2845)

QY 73 LysIleLeuGlnProHisAlaValLeuPheValAspGlu---AspValValGluIle 91
 Db 2 AAAATTCAGAGCAACTCATGCAGTCTCTTTGTTGATGAAAAGGATGTTGTAGAGATA 61

QY 92 AsnGluLysPheThrGluLeuLeuAlaIleThrAsnCysGluGluArgPheSerLeu 111
 Db 62 AATGAAAAGTTCAAGAGTTACTTTTGGCAATTACCAATTGTGAGGAGAGTTTCAGCCTG 121

QY 112 PheLysAsnArgAsnArgLeuSerLysGlyLeuGlnIleAspValGlyCysProValIys 131
 Db 122 TTTAAAACACGAACAGACTAAGTAAGGGCTCTCAATAGACGTGGGTCTCTCTGTGAAA 181

QY 132 ValGlnLeuArgSerGlyGluGluLysPheProGlyValValArgPheArgGlyProLeu 151
 Db 182 GTACAGCTGAGATCTGGGGAGAAAATTTCTCTGGAGTTGTACGCTTCAGAGGACCCCTG 241

QY 152 LeuAlaGluArgThrValSerGlyIlePhePheGlyValGluLeuLeuGluGlyArg 171
 Db 242 TTAGCAGAGAGAGACAGTCTCCGGAAATATCTTTGGAGTTGAAATGCTGGAAGAGGCTGT 301

QY 172 GlyGlnGlyPheThrAspGlyValTyrGlnGlyLysGlnLeuPheGlnCysAspGluAsp 191
 Db 302 GGTCAAGGTTTCACTGACGGGTGTACCAAGGGGAAACAGCTTTTTCAGTGTGATGAAGAT 361

QY 192 CysGly---PheValAlaLeuAspLysLeuGluLeuIleGluAspAspThrAlaLeu 210
 Db 362 TGTGGCGTGTGTTGTCATTTGGCAAGCTAGAACTCATAGAGATGATGACACTGCATTTG 421

QY 211 GluSerAspTyrAlaGlyProGlyAspThrMetGlnValGluLeuProProLeuGluIle 230
 Db 422 GAAAGTATTACGACAGTCTCTGGGACACAATCAGGTGCAACTCTCTCTTTGGAAATA 481

QY 231 AsnSerArgValSerLeuLysGlyGluThrIleGluSerGlyThrValIlePheCys 250
 Db 482 AACTCCAGAGTTTCTTTTGAAGGTGTGAGAAACAATAAGATCTCGAACAGTTATATTTCTGT 541

QY 251 AspValLeuProGlyLysGluSerLeuGlyTyrPheValGlyValAspMetAspAsnPro 270
 Db 542 GATGTTTTGCCAGAAAGAAAGCTTAGGATATTTTGTGTGGTGGACATGGATAACCCCT 601

QY 271 IleGlyAsnTyrAspGlyArgPheAspGlyVal---LeuCysSerPheAlaCysValGlu 289
 Db 602 ATTGGCAACTGGATGGAAGATTTGATGGAGTCAGCTTTGTAGTTTGGTGTGTTGAA 661

QY 290 SerThrIleLeuLeuHisIleAsnAspIleIleProGluSerValThrGlnGluArg 309
 Db 662 AGTACAAATCTATTGTCACATCAATGATATCATCCAGAGAGTGTGACGACGAGGAAGAGG 721

QY 310 ProProLysLeuAlaPheMetSerArgGlyValGlyAspLysGlySerSerSerHisAsn 329
 Db 722 CTTCCCAAACTTGCCCTTTATGTCAAGAGGTGTGGGGACAAAGGTTTCATCCAGTCAAT 781

QY 330 LysProLysAlaThrGlySerThrSerAspProGlyAsnArg---ArgSerGluLeuPhe 348
 Db 782 AAACCAAGGCTACAGAGTCTACTCAGACCTTGAAATAGAACACAGATCTGAATTTATTT 841

QY 349 TyrThrLeuAsnGlySerSerValAspSerGlnProGlnSerLysSerLysAsnThrTyr 368
 Db 842 TATACCTTAAATGGGTCTTCTGTGACTCACAACCAATCCAAATCAAAAATACATGG 901

QY 369 TyrIleAspGluValAlaGluAspProAlaLysSerLeuThrGluIleSerThrAspPhe 388
 Db 902 TACATTGATGAAGTTGCAGAAAGACCTCGCAAAATCTCTTACAGAGATATCTACAGACTTT 961

QY 389 AspArgSerSerProProLeuGlnProProValAsnSerLeuThrThrGluAsnArg 408
 Db 962 GACCGTTCTTACACACACTCCAGCTCTCTCTGTGACTCCTGACACCGAGAACAGA 1021

QY 409 PheHisSerLeuProPheSerLeuThrLysMetProAsnThrAsnGlySerIleGlyHis 428
 Db 1022 TTCCACTCTTTTACCATTCAGTCTCACAAGATGCCCAATACCAATGGAAGTATTGGCCAC 1081

QY 429 SerProLeuSerLeuSerAlaGlnSerValMetGluLeuLeuAsnThrAlaProValGln 448
 Db 1082 AGTCCACTTTCTCTGTGAGCCCACTGTGTAATGGAAGAGCTAAACACTGCACCCGCTCAA 1141

QY 449 GluSerProProLeuAlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeu 468
 Db 1142 GAGAGTCCACCTTGGCCATGCTCTCTGGGAATCATCATGCTCTAGAAGTGGGCTCATTTG 1201

QY 469 AlaGluValLysGluAsnProProPheTyrGlyValIleArgTyrIleGlyGlnProPro 488
 Db 1202 GCTGAAGTTAAGGAGAACCTCTCTTCTATGGGGTAATCCGTTGATCGGTGACCCACCA 1261

QY 489 GlyLeuAsnGluValLeuAlaGlyLeuGluLeuGluAspGluCysAlaGlyCysThrAsp 508
 Db 1262 GGACTGAATGAAGTGTCTGCTGGAGCTGGAACCTGGAAGATGAGTGTGAGGCTGTACCGAT 1321

509 GlyThrPheArgGlyThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPheValLys 528
1322 GGAACTTTCAGAGCACTCGTATTTTCACTGTGCGCTGGAAGAGCGCTTTGTGAAA 1381
529 LeuLysSerCysArgProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIle 548
1382 CTGAAGAGCTGCAGGCTGACTCTAGCTTTCATCATTCGAGCGGTTTCCATCAGATT 1441
549 GluArgCysAsnSerLeuAlaPheGlyTyrLeuSerGluValValGluGluAenThr 568
1442 GAGCGCTGTAACTCTTAGCATTTGGAGGCTACTTAAGTGAAGTAGTAGAAGAAATACT 1501
569 ProProLysMetGluLysGluGlyLeuGluIleMetIleGlyLysLysLysGlyIleGln 588
1502 CCACCAAAATGGAAGAAGGCTTGGAGATAATGATTGGGAAGAAAGGACATCCAG 1561
589 GlyHisTyrAsnSerCysTyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSer 608
1562 GGTCAATTAACATCTTGTACTTAGACTCAACCTTATCTGCTTATTTGCTTTAGTTCT 1621
609 ValLeuAspThrValLeuLeuArgProLysGluLysAsnAspValGluTyrTyrSerGlu 628
1622 GTTCTGGACACTGTGTACTTGGACCCAAAGAAAGACGATGAGATATATTATAGTAA 1681
629 ThrGlnGluLeuLeuArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCys 648
1682 ACCCAAGAGCTACTGAGGACAGAAATGTAAATCTCTGAGATATATGATATGTGTGT 1741
649 AlaThrLysIleMetLysLeuArgLysIleLeuGluLysValGluAlaLaserGlyPhe 668
1742 GCACAAAATATGAACCTGAGAAATATCTTGAAGAGTGAGGCTGCATCAGGATTT 1801
669 ThrSerGluGluLysAspProGluGluPheLeuAsnIleLeuPheHisIleLeuArg 688
1802 ACCTCTGAAGAAAAGATCTGAGGATCTTGAATATCTGTTTCAATATTTAAGG 1861
689 ValGluProLeuLysLysIleArgSerAlaGlyLysValGlnAspCysTyrPheTyr 708
1862 GTAGAACCCTTGTAAAAATAAGATCAGCAGGTCAAAAGGTCAAGATGTTACTTCTAT 1921
709 GlnIlePheMetGluLysAsnGluLysValGlyValProThrIleGlnLeuLeuGlu 728
1922 CAAATTTTATGGAATAAATGAGAAGTTGGCGTTCCCAATCAGCAGTTGTTAGAA 1981
729 TrpSerPheIleAsnSerAsnLeuLysPheAlaGluAlaProSerCysLeuIleIleGln 748
1982 TGGTCTTTTATCAACAGTAACCTGAATTTGAGAGCCATCATGTCTGATTATTCAG 2041
749 MetProArgPheGlyLysAspPheLysLeuPheLysLysIlePheProSerLeuGluLeu 768
2042 ATGCTCTGATTTGGAAGAGACTTAAACTATTTTAAAAAATTTTCTCTCTGGAATTA 2101
769 AsnIleThrAspLeuLeuGluAspThrProArgGlnCysArgIleCysGlyLysLeuAla 788
2102 AATATAACAGATTTACTTGAAGACACTCCAGACAGTGGCGGATATGTGGAGGCTTGCA 2161
789 MetTyrGluCysArgGluCysTyrAspAspProAspIleSerAlaGlyLysIleLysGln 808
2162 ATGTATGAGTGTAGAGAAATGCTACGACGATCCGACATCTCAGCTGGAATAATCAAGCAG 2221
809 PheCysLysThrCysAsnThrGlnValHisLeuHisProLysArgLeuAsnHisLysTyr 828
2222 TTTTGTAAACCTGCACACTCAAGTCCACCTTCATCCGAGAGGCTGAATCATTAATAT 2281
829 AsnProValSerLeuProLysAspLeuProAspTrpAspTrpArgHisGlyCysIlePro 848
2282 AACCCAGTGTCACTTCCCAAGACTTACCCGACTGGGACTGGAGACACGCGCTCATCCCT 2341
849 CysGlnAsnMetGluLeuPheAlaValLeuCysIleGluThrSerHisTyrValAlaPhe 868
2342 TGCCAGAAATAGGAGTATTTGCTGTCTCTGATAGAAACAAAGCCCATATGTTGCTTTT 2401
869 ValLysTyrGlyLysAspAspSerAlaTrpLeuPhePheAspSerMetAlaAspArgAsp 888

2402 GTGAAGTATGGGAAGGACGATTCTGCTGCTCTTCTTTGACAGATGGCGATGGGAT 2461
889 GlyGlyGlnAsnGlyPheAsnIleProGlnValThrProCysProGluValGlyGluTyr 908
2462 GGTGGTCAGAAATGGCTTCAACATCTCTCAAGTCAACCCATGCCCAGAGTAGGAGATAC 2521
909 LeuLysMetSerLeuGluAspLeuHisSerLeuAspSerArgArgIleGlnGlyCysAla 928
2522 TTGAAGATGTCTCTGGAAGACCTGCAATCTCTTGGACTCCAGGAGAAATCCAAGGCTGTGCA 2581
929 ArgArgLeuLeuCysAspAlaTyrMetCysMetTyrGlnSerProThrMetSerLeuTyr 948
2582 CGAAGACTGCTTTGTGATCATATATGATGATGATGATGATGATGATGATGATGATGAT 2641
949 Lys 949
2642 AAA 2644

RESULT 13
AX883937
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
CDS

AX883937
Sequence 18842 from Patent EP1074617.
AX883937
AX883937.1 GI:40038838
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.
Primers for synthesising full-length cDNA and their use
Patent: EP 1074617-A 18842 07-FEB-2001; *data*
Research Association for Biotechnology (JP)

1. .2845
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
452. .2647
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAE92170.1"
/db_xref="GI:40038839"
/translation="MQVELPPLINRSVSKVGETIESGTIVICDVLPGKESLGYFVG
VMDNPIGNWDGFRDGVQLCSFACVESTILLHNDIIPESVTQERRPPKLAIFMRGVG
DKSSSHNPKPKATGSTDPGNRNSLFYTLNGSSVDSQPSQSKNTWYIDEVADPA
KSLTEITDPDRSSPPLQPPVNSLTENRPHSLPFLSLTKMPTNGSIGHSPGLSQA
SVMEELNTPAQVQSPPLAMPNGSHGLVGLAEVKENPPFYGVIRWIGQPPGLNEVL
AGLELEDCAGCTDGTFRGTRVFTCALKKALFKVLSKSPDRFASLQPVSNQIERCN
SLAFGYLSEVEENTPPRMEKEGLEIMIGKKGIQGHYNSCYLDLSTFLCFLFAFSSVL
DVLVLPKKNDEYVYSETQELLRLTEIVNPLRIYGVYCATKIMKRLKILEKVEASGF
TSEKDPPEFLNLFPHILRLVEPLKIRSAQVQDCFYQIEMNEKKGVPVPTIQQL
LEWFSFNLSLAFAPSLIIOMPRFKGDFLKKIIPPSLEINITDLEEDTPROCRIC
GGLAMTECRECYDDPDISAGIKQFCNTQVHLHPKLNHKNPVSLPKDLDPDW
RHGCIQCNWELPVLICIESHYVAFTVKGDSDAWLFFDSMDRGGQNGFNIPQVT
PCPEVGEYLKMSLEDLHSLDSLRRIQGCARRLLCDAYMCMQSTNLSLYK"

Alignment Scores:
Pred. No.: 0 Length: 2845
Score: 4601.00 Matches: 875
Percent Similarity: 99.32% Conservative: 0
Best Local Similarity: 99.32% Mismatches: 2
Query Match: 91.40% Indels: 4
DB: 6 Gaps: 4

US-09-671-687A-3 (1-949) x AX883937 (1-2845)

73 LysIleLeuGluGlnProHisAlaValLeuPheValAspGlu---AspValValGluIle 91

Db 2 AAAATCTTAGAGCAACCTCATGCAGTCTCTTTGTTGATGAAAGGATGTTGTAGAGATA 61
Qy 92 AnGluLyPheThrGluLeuLeuAlaIleThrAsnCysGluGluArgPheSerLeu 111
Db 62 AATGAAAAGTTTACAGAGTTACTTTTGGCAATTAACCAATTTGTGAGGAGGTTTCAGCCTG 121
Qy 112 PheLyAsnArgAsnArgLeuSerLyGlyLeuGlnIleAspValGlyCysProValLyS 131
Db 122 TTTTAAAAACAGAAACAGACTAAGTAAGGCTCCAAATAGACGTGGGCTGTCTGTGAA 181
Qy 132 ValGlnLeuArgSerGlyGluGluLyS PheProGlyValValArgPheArgGlyProLeu 151
Db 182 GTACAGCTGAGATCTGGGAGAGAAAATTTCTGGAGTTGTACGCTTCAGAGACCCCTG 241
Qy 152 LeuAlaGluArgThrValSerGlyIlePhePheGlyValGluLeuLeuGluGluArg 171
Db 242 TTAGCAGAGAGACAGTCTCCGGAATATTTCTTTGGAGTTGAAATGCTGGAAGAAGTCTGT 301
Qy 172 GlyGlnGlyPheThrAspGlyValTyrGlnGlyLySLeuPheGlnLeuPheGlnCysAspGluAsp 191
Db 302 GGTCAAGGTTTCACTGACGGGGTGTACCAAGGGAAACAGCTTTTTCAGTGTGATGAAGAT 361
Qy 192 CysGly- --PheValAlaLeuAspLySLeuGluLeuIleGluAspAspThrAlaLeu 210
Db 362 TGTGGCGTGTGTTGTCATTTGGACAAGCTAGAACTCATGAAGATGATGACACTGCAATG 421
Qy 211 GluSerAspTyrAlaGlyProGlyAspThrMetGlnValGluLeuProProLeuGluIle 230
Db 422 GAAAGTATTACGAGTCTCTGGGACACATGACAGTGCNACTTCTCTTTGGAAATA 481
Qy 231 AsnSerArgValSerLeuLyS GlyGlyGluThrIleGluSerGlyThrValIlePheCys 250
Db 482 AACTCCAGAGTTTCTTTGAAGTTGGAGAAACAATAGAACTCTGGAACAGTTATATTCTGT 541
Qy 251 AspValLeuProGlyLyS GluSerLeuGlyTyrPheValGlyValAspMetAspAsnPro 270
Db 542 GATGTTTTCCAGGAAAGAAAGCTTAGGATATTTTGGTGTGACATGGATAACCCCT 601
Qy 271 IleGlyAsnTyrAspGlyArgPheAspGlyVal- --LeuCysSerPheAlaCysValGlu 289
Db 602 ATTGGCAACTGGGATGGAGATTGATGAGTGACGCTTTGTAGTTTGGGTGTGTGAA 661
Qy 290 SerThrIleLeuLeuHisIleAsnAspIleIleProGluSerValThrGlnGluArgArg 309
Db 662 AGTACAATTTCTATTGCATCAATGATATCATCCAGAGAGTGTGCGAGGAAAGAGG 721
Qy 310 ProProLySLeuAlaPheMetSerArgGlyValGlyAspLyS GlySerSerSerHisAsn 329
Db 722 CCTCCAAACTTGGCTTTATGTCAAGAGGTGTTGGGGACAAAGGTTTCATCCAGTCAAT 781
Qy 330 LysProLySAlaThrGlySerThrSerAspProGlyAsnArg- --ArgSerGluLeuPhe 348
Db 782 AAACCAAGGCTACAGATCTACTCAGACCTCGGAAATAGAAACAGATCTGAATATTAT 841
Qy 349 TyrThrLeuAsnGlySerValAspSerGlnProGlnSerLyS SerLySAsnThrTyr 368
Db 842 TATACCTTAATGGGCTTCTGTTGACTCACACCAATCCAAATCAAAAATACATGG 901
Qy 369 TyrIleAspGluValAlaGluAspProAlaLyS SerLeuThrGluIleSerThrAspPhe 388
Db 902 TACATTGATGAAGTTGCAGAAAGCCCTGCAAAATCTCTTACAGAGATATCTACAGACTTT 961
Qy 389 AspArgSerSerProLeuGlnProProValAsnSerLeuThrThrGluAsnArg 408
Db 962 GACCGTCTTCCACCACTCCAGCTCTCTCTGTAACCTCACTGACCACCCGAGAACAGA 1021
Qy 409 PheHisSerLeuProPheSerLeuThrLyS MetProAsnThrAsnGlySerIleGlyHis 428
Db 1022 TTCCACTCTTACCATTCAGTCTCACCAGATGCCCCAATACCAATGGAAGTATGGCCAC 1081
Qy 429 SerProLeuSerLeuSerAlaGlnSerValMetGluLeuLeuAsnThrAlaProValGln 448
Db 1082 AGTCCACTTTCTCTGTGAGCCAGTCTGTAATGGAAGAGCTAAACACTGCAACCGTCCAA 1141

Qy 449 GluSerProProLeuAlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeu 468
Db 1142 GAGAGTCCACCTTGGCCATGCTCTCTGGAACTCACAATGGTCTAGAAAGTGGGCTCATG 1201
Qy 469 AlaGluValLyS GluAsnProProPheTyrGlyValIleArgTrpIleGlyGlnProPro 488
Db 1202 GCTGAAGTTTAAAGAGAACCTCTCTTCTATGGGGTAATCCGTTGGATCGGTGACGCCA 1261
Qy 489 GlyLeuAsnGluValLeuAlaGlyLeuGluLeuGluAspGluCysAlaGlyCysThrAsp 508
Db 1262 GGACTGAAATGAAGTGTCTGCTGAGCTGGAACTGGAAAGATGAGTGTGCAGGCTGTACCGAT 1321
Qy 509 GlyThrPheArgGlyThrArgTyrPheThrCysAlaLeuLyS LySAlaLeuPheValLyS 528
Db 1322 GGAACCTTCAGAGGCACTCGGTATTTTCCCTGTGCCCTGAAGAAGGCGCTGTTTGTGAA 1381
Qy 529 LeuLyS SerCysArgProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIle 548
Db 1382 CTGAAGAGCTGACGGCTGACTCTAGGTTTGCATCATTTGAGCGCGGTTTCCAATCAGATT 1441
Qy 549 GluArgCysAsnSerLeuAlaPheGlyGlyTyrLeuSerGluValValGluGluAsnThr 568
Db 1442 GAGCGCTGTAACTCTTTAGCAITTTGAGGCTACTTAAGTGAAGTAGTAGAAGAAATACT 1501
Qy 569 ProProLyS MetGluLyS GluGlyLeuGluIleMetIleGlyLyS LyS LyS GlyIleGln 588
Db 1502 CCACCAAAATGGAAAGAGGCTTGGAGATAATGATTGGGAAGAAAGGCATCCAG 1561
Qy 589 GlyHisTyrAsnSerCysTyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSer 608
Db 1562 GGTCAATCAAAATTTCTTTACTTAGACTCAACCTTATTTCTGCTTATTTGCTTTTGTCT 1621
Qy 609 ValLeuAspThrValLeuLeuArgProLyS GluLyS AsnAspValGluTyrTyrSerGlu 628
Db 1622 GTTCTGGACACTGTGTACTTGGACCAAGAAAGAACGATGTAGAATATTATTAGTGAA 1681
Qy 629 ThrGlnGluLeuLeuArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCys 648
Db 1682 ACCAAGAGCTACTGAGGACAGAAATTTGTAATCTCTGAGATAATATGATATGTCTGT 1741
Qy 649 AlaThrLyS IleMetLyS LeuArgLyS IleLeuGluLyS ValGluAlaAlaSerGlyPhe 668
Db 1742 GCCACAAAATTTATGAACTTGAGAAAATACTTTGAAAGGTGGAGGCTGCATCAGGATT 1801
Qy 669 ThrSerGluGluLyS AspProGluGluPheLeuAsnIleLeuPheHisIleLeuArg 688
Db 1802 ACCTCTGAAGAAAGATCCTGAGGAATTTCTGAATATTCTGTTTCATCATATTTTAAGG 1861
Qy 689 ValGluProLeuLeuLyS IleArgSerAlaGlyGlnLyS ValGlnAspCysTyrPheTyr 708
Db 1862 GTAGAACCCTTTGTAAATAATAGATCAGCAGGTCMAAGGTACAGATTGTTACTTCTAT 1921
Qy 709 GlnIlePheMetGluLyS AsnGluLyS ValGlyValProThrIleGlnGlnLeuLeuGlu 728
Db 1922 CAAATTTTATGAAAAAATAGAAAGTTGGCGTTCCCAAAATTCAGCAGTTGTTAGAA 1981
Qy 729 TrpSerPheIleAsnSerAsnLeuLyS PheAlaGluAlaProSerCysLeuIleIleGln 748
Db 1982 TGGTCTTTTATCAACAGTAACTCGAAATTTGCAAGGCAACCATCATCTCTGATTATTTCAG 2041
Qy 749 MetProArgPheGlyLyS AspPheLyS LeuPheLyS LyS IlePheProSerLeuGluLeu 768
Db 2042 ATGCTTCGATTGGAAGAGACTTTAAACTATTTTAAAAAATTTTCTCTCTCGAATTA 2101
Qy 769 AsnIleThrAspLeuLeuGluAspThrProArgGlnCysArgIleCysGlyLeuAla 788
Db 2102 AATATAACAGATTTACTTGAAGACACTCCCAAGCAGTGGCGGATATGTGGAGGCTTGCA 2161
Qy 789 MetTyrGluCysArgGluCysTyrAspAspProAspIleSerAlaGlyLyS IleLyS Gln 808
Db 2162 ATGTATGAGTGTAGAAATGCTACGAGTCCGACATCTCAGCTGGAANAATCAAGCAG 2221

Qy	192	CysGly---PheValAlaLeuAspLysLeuGluLeuLeuGluAspAspThrAlaLeu	210	Db	1442	GAGCGCTGTAACCTTTAGCATTTGAGGCTACTTAAAGTGAAGTAGTAGAAGAAATACT	1501
Db	362	TGTGGCGTGTGTGTGCAATGGACAAGCTAGAACTCATAGAAGATGATGACACTGCATG	421	Qy	569	ProProLysMetGluLysGluGlyLeuGluLeuMetIleGlyLysLysLysGlyIleGln	588
Qy	211	GluSerAspTyrAlaGlyProGlyAspThrMetGlnValGluLeuProLeuGluIle	230	Db	1502	CCACCACAAANTGGAAAAGAGGCTTGGAGATAATGATTTGGGAAGAAAGAGCANTCCAG	1561
Db	422	GAAAGTGATTAACGACAGGCTCGGGACACAATGCAAGGTGCAACTTCCTCTTTGGAAATA	481	Qy	589	GlyHisTyrAsnSerCysTyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSer	608
Qy	231	AsnSerArgValSerLeuLysGlyGlyGluThrIleGluSerGlyThrValIlePheCys	250	Db	1562	GGTCATTACAATTTCTTTACTTTAGACTCAACCTTATCTGCTTATTTGCTTTTAGTTCT	1621
Db	482	AACTCCAGAGTTCTTTGAGGTGGAGNAACAATAGATCTGGAAACAGTTATTTCTGT	541	Qy	609	ValLeuAspThrValLeuLeuArgProLysGluLysAsnAspValGluTyrTyrSerGlu	628
Qy	251	AspValLeuProGlyLysGluSerLeuGlyTyrPheValGlyValAspMetAspAsnPro	270	Db	1622	GTTCTGGACACTGTGTACTTGGACCCAAAGAAAGACGATAGAAATATTAGTGAA	1681
Db	542	GATGTTTTGCCAGGAAAAGAAAGCTTAGCATATTTGTTGGTGTGGACATGATCAACCT	601	Qy	629	ThrGlnGluLeuLeuArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCys	648
Qy	271	IleGlyAsnTrpAspGlyArgPheAspGlyVal---LeuCysSerPheAlaCysValGlu	289	Db	1682	ACCAAGAGCTACTGAGGACAGAAATTTGTAATCTCTGAGAAATATATGATATGTGT	1741
Db	602	ATTGGCAACTGGGATGGAGATTTGATGGAGTGCAGCTTTGAGTTTTCGCTGTGTGAA	661	Qy	649	AlaThrLysIleMetLysLeuArgLysIleLeuGluLysValGluAlaAlaSerGlyPhe	668
Qy	290	SerThrIleLeuLeuHisIleAsnAspIleIleProGluSerValThrGlnGluArgArg	309	Db	1742	GCCACAAATTTATGAACTGAGGAAATATTTGAAAGGTGAGGCTGCATCAGGATTT	1801
Db	662	AGTACAATTTATGACATCAATGATATCATCCACAGAGGTGACGAGGAAGAGG	721	Qy	669	ThrSerGluGluLysAspProGluGluPheLeuAsnIleLeuPheHisIleLeuArg	688
Qy	310	ProProLysLeuAlaPheMetSerArgGlyValGlyAspLysGlySerSerSerHisAsn	329	Db	1802	ACCTCTGAAGAAAAAGATCCTGAGGAATTTCTTGAATATTTCTGTTTCATCATATTTTAAG	1861
Db	722	CCTCCCAACTTGCCTTTATGTCAAGAGGTGTGGGACAAAGGTTTCATCCAGTCATAT	781	Qy	689	ValGluProLeuLeuLysIleArgSerAlaGlyGlnLysValGlnAspCysTyrPheTyr	708
Qy	330	LysProLysAlaThrGlySerThrSerAspProGlyAsnArg---ArgSerGluLeuPhe	348	Db	1862	GTAGAACCTTTGCTAAAAATAAGATCAGCAGGTCAAAAGGTACAAGATTTGATCTCTAT	1921
Db	782	AAACCAAGGCTACAGGATCTACCTCAGACCTCGGAATAGAAACAGATCTGAATTTAT	841	Qy	709	GlnIlePheMetGluLysAsnGluLysValGlyValProThrIleGlnGlnLeuLeuGlu	728
Qy	349	TyrThrLeuAsnGlySerSerValAspSerGlnProGlnSerLysSerLysAsnThrTrp	368	Db	1922	CAATTTTATGGAAAAAATGAAAGTTGGGTTCCCAATTTAGCAGTTGTTAGAA	1981
Db	842	TATACCTTAATTTGGGTCTCTGTGTGACTCACACCAATCCCAATCCAAATAATACAT	901	Qy	729	TrpSerPheIleAsnSerAsnLeuLysPheAlaGluAlaProSerCysLeuIleIleGln	748
Qy	369	TyrIleAspGluValAlaGluAspProAlaLysSerLeuThrGluIleSerThrAspPhe	388	Db	1982	TGCTCTTTTATCAACAGTAACCTGAAATTTGCAGAGCACCACATCTGCTGATTTATCAG	2041
Db	902	TACATTTGATGAAGTTGCAGAGACCTCGCAAAATCTCTTACAGAGATATCTACAGACT	961	Qy	749	MetProArgPheGlyLysAspPheLysLeuPheLysLysIlePheProSerLeuLeuLeu	768
Qy	389	AspArgSerSerProProLeuGlnProProValAsnSerLeuThrThrGluAsnArg	408	Db	2042	ATGCTCTGATTTGGAAAAAGACTTTAAACTATTTTAAAAAAATTTTCTCTCTGGAATTA	2101
Db	962	GACCGTCTTCCACCACCTCCAGCCTCTCTCTGTGAATCACTGACACCCAGGAGACAGA	1021	Qy	769	AsnIleThrAspLeuLeuGluAspThrProArgGlnCysArgIleCysGlyGlyLeuAla	788
Qy	409	PheHisSerLeuProPheSerLeuThrLysMetProAsnThrAsnGlySerIleGlyHis	428	Db	2102	ATATTAACAGATTTACTTGAAGACACTCCACGACAGTCCGGATATGTGGAGGCTTGCA	2161
Db	1022	TTCCACTCTTTACCATTTAGTCTCACCAAGATGCCCAATACCAATGGAAGTATTTGCC	1081	Qy	789	MetTyrGluCysArgGluCysTyrAspAspProAspIleSerAlaGlyLysIleLysGln	808
Qy	429	SerProLeuSerLeuSerAlaGlnSerValMetGluGluLeuAsnThrAlaProValGln	448	Db	2162	ATGTATGATGTAGAGATGCTACGACGATCCGACATCTCAGCTGGAATAAATCAAGCAG	2221
Db	1082	AGTCCACTTTCTGTGAGCCAGTCTGTAATGGAAGAGCTAAACACTGCACCCGTCCAA	1141	Qy	809	PheCysLysThrCysAsnThrGlnValHisLeuHisProLysArgLeuAsnHisLysTyr	828
Qy	449	GluSerProProLeuAlaMetProProGlyAsnSerHisGlyLeuValGlySerLeu	468	Db	2222	TTTTGTAAACCTTCGCAACACTCAAGTCCACCTTATCCCGAAGAGGCTGAATCATATAAT	2281
Db	1142	GAGAGTCCACCTTTGGCCATGCTCTCTGGAACTCACTGCTAGAGGTGGGCTCATTTG	1201	Qy	829	AsnProValSerLeuProLysAspLeuProAspTrpAspTrpArgHisGlyCysIlePro	848
Qy	469	AlaGluValLysGluAsnProProPheTyrGlyValIleArgTrpIleGlnProPro	488	Db	2282	AACTCCAGTGTCACTTCCCAAGACTTTACCCGACTGGGACTTGGAGACACGGCTGATCCCT	2341
Db	1202	GCTGAAGTTAAGGAGAACCTCTCTTTCTATGGGGTAAATCCGTTGGATCGGTACGCCA	1261	Qy	849	CysGlnAsnMetGluLeuPheAlaValLeuCysIleGluThrSerHisTyrValAlaPhe	868
Qy	489	GlyLeuAsnGluValLeuAlaGlyLeuGluLeuGluAspGluCysAlaGlyCysThrAsp	508	Db	2342	TGCCAGATATGAGTTATTTGCTGTTCTCTGCATAGAAACAAGCCACTATGTTGCTTTT	2401
Db	1262	GGACTGAATGAAGTCTCGCTGGACTGGAACTGGAAGATGAGTGTGCGAGCTGTACGGAT	1321	Qy	869	ValLysTyrGlyLysAspAspSerAlaTrpIlePhePheAspSerMetAlaAspArgAsp	888
Qy	509	GlyThrPheArgGlyThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPheValLys	528	Db	2402	GTGAAGTATGGGAAGGACGATTTCTGCTGGCTCTTCTTTCGACAGCATGGCCGATCGGAT	2461
Db	1322	GGAACTTCAGAGGCACTCGTATTTTCACTGTGCCCTGAAGAGGCGCTGTTGTGAA	1381	Qy	889	GlyGlyGlnAsnGlyPheAsnIleProGlnValThrProCysProGluValGlyGluTyr	908
Qy	529	LeuLysSerCysArgProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIle	548	Db	2462	GGTGGTTCAGAAATGGCTTCAACATTCCTCAAGTCAACCCATCCCAAGAGTAGAGAGTAC	2521
Db	1382	CTGAAGAGCTGACGCGCTGACTAGGTTTGCATCATTTGACGCGGTTTCCAAATCAGATT	1441	Qy	909	LeuLysMetSerLeuGluAspLeuHisSerLeuAspSerArgArgIleGlnGlyCysAla	928
Qy	549	GluArgCysAsnSerLeuAlaPheGlyGlyTyrLeuSerGluValValGluGluAsnThr	568				

Db 2522 TTGAAGATGCTCTCTGGAAGACCTGCGATTCTTGGACTCCAGGAGAAATCCAAAGCTGTGCA 2581

QY 929 ArgA9LeuLeuCuSAspAlaTyrMetCysMetTyrGlnSerProThrMetSerLeuTyr 948
 |||||

Db 2582 CGAAGACTCTTTGTGATGCATATATGTGCTACCATGTCACAGATCCAAATGAGTTTGATC 2641
 |||||

QY 949 Lys 949

Db 2642 AAA 2644
 |||||

RESULT 15
 AC123449

LOCUS Rattus norvegicus clone CH230-21D6, *** SEQUENCING IN PROGRESS ***,
 3 unordered pieces.

DEFINITION AC123449

ACCESSION AC123449

VERSION HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.

KEYWORDS Rattus norvegicus (Norway rat)

SOURCE Rattus norvegicus

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE 1 (bases 1 to 241990)

AUTHORS Muzny,D,Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
 Allen,C., Aillen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
 Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
 Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
 Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
 Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
 Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
 Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
 Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
 Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
 Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
 Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
 Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
 Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
 Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
 Gregorogis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
 Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hernandez,J.,
 Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
 Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogue,M.,
 Hollins,B., Howells,S., Huiyik,S., Hume,J., Idlebird,D., Jackson,A.,
 Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
 Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
 Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
 Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
 Lorensuhera,L., Loulseghe,H., Lozado,R.J., Lu,X., Ma,J.,
 Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
 Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
 Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Moore,S.,
 Milosavljevic,A., Miner,G., Mirja,E., Montemayor,J., Moore,S.,
 Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
 Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
 Nwaokeme,O., Okwuonu,G., Olarnpusagoon,A., Pal,S., Parks,K.,
 Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,
 Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
 Puzo,M., Quirroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
 Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
 Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
 Sanders,W., Savary,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
 Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
 Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
 Steimle,M., Strong,R., Sutton,A., Svatek,A., Taber,P., Taylor,C.,
 Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
 Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
 Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,P.,
 Williams,G., Willson,R., Wlezyk,R., Wooden,H., Worley,K.,
 Wright,D., Wright,J., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
 Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
 Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
 Weinstein,G. and Gibbs,R.A.

TITLE Direct Submission

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 241990)

AUTHORS Worley,K.C.

TITLE Direct Submission

JOURNAL Submitted (29-MAY-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

REFERENCE 3 (bases 1 to 241990)

AUTHORS Rat Genome Sequencing Consortium.

TITLE Direct Submission

JOURNAL Submitted (21-SEP-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

COMMENT On Sep 21, 2002 this sequence version replaced gi:21902787.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (http://www.hgsc.bcm.tmc.edu/projects/rat/). As a result, the
 sequence may extend beyond the ends of the clone and there may be
 contigs that consist entirely of whole genome shotgun sequence
 reads. Both end sequences and whole genome shotgun sequence only
 contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GXOF

Center clone name: CH230-21D6

----- Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 229402 bases at least Q40

Consensus quality: 231763 bases at least Q30

Consensus quality: 233057 bases at least Q20

Estimated insert size: 260790; sum-of-contigs estimation

Quality coverage: 3x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)

* NOTE: This sequence may represent more than one 'clone'.
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 3 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 238985: contig of 238985 bp in length
 * 238986 239085: gap of unknown length
 * 239086 240731: contig of 1646 bp in length
 * 240732 240831: gap of unknown length
 * 240832 241990: contig of 1159 bp in length.

FEATURES

source

1. .241990
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-21D6"
 1323. .2205
 /notes="clone boundary
 clone end:Sp6
 site:ECORI
 end_sequence:BH269859"
 complement(236297. .237124)
 /notes="clone boundary
 clone end:T7
 site:ECORI
 end_sequence:BH269858"

misc_feature

misc_feature

misc_feature

ORIGIN

Alignment Scores:

Pred. No.:	0	Length:	241990		
Score:	4600.50	Matches:	883		
Percent Similarity:	95.29%	Conservative:	27		
Best Local Similarity:	92.46%	Mismatches:	38		
Query Match:	91.39%	Indels:	8		
DB:	2	Gaps:	5		
US-09-671-687A-3 (1-949) x AC123449 (1-241990)					
Qy	1	MetSerSerGlyLeuTyrSerGlnGluLysValThrSerProTyrTyrGluGluArgIle	20		
Db	205375	ATGAGTTTCAGGCTGTGGAAACCAAGAGAAAGTTACTTTCACCTTACCTTGGGAAGAACGGCTT	205434		
Qy	21	PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLeuLys	40		
Db	205435	TTTTATCTCTCTTCAAGATGCAAGTAAAC---AAACAGACACAGAAAGCTCTCTGAGA	205490		
Qy	41	ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisArgIle	60		
Db	205491	GTACCCAAAGGAGCATAGACAGTACATCCAAAGACTGTTCCGTGGGGCATTCACAGATT	205550		
Qy	61	ProSerAlaLysGlyLysLysGlnIleGlyLeuLysIleLeuGlnProHisAla	80		
Db	205551	CCTTCTGCTTAAAGGCAAGAAATACAGATTGGATTAAATAATCTTAGCAACACCGCATGCA	205610		
Qy	81	ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu	99		
Db	205611	GTTCTGTTGTGTGATAAAGGATGTTGTAGAAATAAATAAATAATTCACAGAGTTACTG	205670		
Qy	100	LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer	119		
Db	205671	TTGGCAATTACCAACTGTGAGGAGAGGCTCAGCCTATTTAGAAACAGAAATCCGACTAAGT	205730		
Qy	120	LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu	139		
Db	205731	AAAGGCTCCAGGTAGACGTGGGCAGTCCCTGTGAGAGTACAGCTGGCATCTGGGGAGGAG	205790		
Qy	140	LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly	159		
Db	205791	AGTTTCCAGGAGTTGTAGCTTTCAGAGGACCTTTATTACGGAGAGGACGGTGTCCGGG	205850		
Qy	160	IlePhePheGlyValGluLeuLeuGluGluGlyArgGlyGlnGlyPheThrAspGlyVal	179		
Db	205851	ATTTCTTTGGAGTAGAATTACTTGGAAAGGTCGTGGCCCAAGTTTCACTGATGGGTG	205910		
Qy	180	TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp	198		
Db	205911	TATCAAGGAAACAGCTCTTCCAGTGTGATGAGGACTGTGGTGTGTTTGTGCAATTGGAC	205970		
Qy	199	LysLeuLeuLeuIleGluAspAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly	218		
Db	205971	AACTCGAGCTTATAGAAGATGATGACAAATGGGTTGGAAGTGAATTTTGCAGGCCACGGA	206030		
Qy	219	AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly	238		
Db	206031	GATACAGTCCAGGTTGAACCTCCCTTTGGAAATAAATCTCCAGAGTTTCTTTGAAGGTT	206090		
Qy	239	GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer	258		
Db	206091	GGAGAAAGTACAGAACTCGAAACAGTGATATCTGTGATGTTTTACCAGGAAAGAGAGT	206150		
Qy	259	LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe	278		
Db	206151	CTAGGATATTTTGTGTGTGGACATGGATAACCCCTATTGGCAACTGGGATGGAAGGTTT	206210		
Qy	279	AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn	297		
Db	206211	GATGGGGTTTCAGGCTTGTCAAGTTTGCAGGTGTTGAGAGTACAGTTCTTCTACACATCAAT	206270		
Qy	298	AspIleIleProGluSerValThrGlnGluArgProProLysLeuAlaPheMetSer	317		
Db	206271	GACATCATCCAGATAGCGTGCACAGAAAGGAGACCTCCCAAACTTGCCTTTATGTCA	206330		
Qy	318	ArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThrGlySerThr	337		
Db	206331	AGAGGTGTAGGTGACAAAGGCTCATCTAGTCATAATAAAACCAAGGTTTACAGGATCTACC	206390		
Qy	338	SerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVal	356		
Db	206391	TCAGACCTTGGAGTAGAAGACAGATCTGAATATTTTATACCTTAAATGGGTCTATCTGTT	206450		
Qy	357	AspSerGlnProGlnSerLysSerLysAsnThrTyrTyrIleAspGluValAlaGluAsp	376		
Db	206451	GACTCACACAACATCCAAAGTCCAAAACCCATGGTACATTGATGAAGTTGCAGAAGAC	206510		
Qy	377	ProAlaLysSerLeuThrGluLysSerThrAspPheAspArgSerSerProProLeuGln	396		
Db	206511	CCTGCAAGTCACTTACAGAGATGTTCTCAGACTCGGACATTCATCGCTCCACCCCAA	206570		
Qy	397	ProProProValAsnSerLeuThrGluAsnArgPheHisSerLeuProPheSerLeu	416		
Db	206571	CCTCCTTCCATGAACCTCTTGTCTAGCGAGAACAGATTCCACTCCTTACCTTCAGCCTG	206630		
Qy	417	ThrLysMetProAsnThrAsnGlySerIleGlyHisSerPro---LeuSerLeuSerAlaG	436		
Db	206631	ACAAAGATGCCCAACACTAATGTCAGCATGGCTCACAGTCCACTCTCTCTCTGTGAGTGC	206690		
Qy	436	InsSerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetP	456		
Db	206691	AGTCTGTGATGGGGAGCTGAACAGCAGCGCTGTCCAGGAGAGTCCACCTTGGCCAGCT	206750		
Qy	456	roProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProp	476		
Db	206751	CTTCTGGGAATGACACCGGCTAGAGGTGGCTCACTGGCTGAAGTAAAGAGAACCCCC	206810		
Qy	476	roPheTyrGlyValIleArgTropIleGlyGlnProProGlyLeuAsnGluValLeuAlaG	496		
Db	206811	CGTCTATG3GGTATTCGGTGTGGATTGGCCAGCCAGGAGTCAAGTGTCTGCTGCTG	206870		
Qy	496	lyLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgT	516		
Db	206871	GATTGGAAACCGGAAGATGAATGCGCAGGTGTCAGCGTGAACCTTTTCAGGGCACCGCT	206930		
Qy	516	yrPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspS	536		
Db	206931	ATTTCACTGTGCTTGAAGAAAGCAGTGTTCGTGAAACTGAAGAGTGCAGACAGACT	206990		
Qy	536	erArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaP	556		
Db	206991	CTAGGTTTGCATCTTGCAGCCTGTTTCCAATCAGATCGAAAGGTGAACCTTTTAGCAT	207050		
Qy	556	heGlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluG	576		
Db	207051	TTGGGGGCTACTTAAAGTAGTAGTAGAAGAAATAACGCCACCTTAAATGGAAAGGAAG	207110		
Qy	576	lyLeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrL	596		
Db	207111	GTTTAGAGATAATGATTGGAAGAAAGAGCATCCAGGGCCATTACAAATTTCTTGTACT	207170		
Qy	596	euAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuA	616		
Db	207171	TAGACTCAACTTTATCTGCTTATTTGCTTTTGTCTGCTGCGCTGGACACTGTATTA	207230		
Qy	616	rgProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrG	636		
Db	207231	GACCCAAAGAGAGAAATGAGTAGAGTATTACAGTAGAGACTCAAGAGCTACTGAGGACG	207290		
Qy	636	luIleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuA	656		
Db	207291	AGATAGTCAATCTCTGAGATATATGATATGTTGTGTGCCACCAAGATTATGAAGCTGA	207350		
Qy	656	rglyIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProG	676		
Db	207351	GGAAAAATCTTGAAGAAAGTTGAGGCTCATCAGGATTATTCCTCTGAGGAAAAAGATCCTG	207410		
Qy	676	luGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleA	696		


```
Db 207411 AGAATTTCTAAACATCCTGTTTCATGATATTTTAAAGGATTGAACCATTTGTTAAAAATAT 207470
QY 696 rGserAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnG 716
Db 207471 GGTCAAGAGTCAAAAAGTTCAAGACTGTAACTTCTATCAAAATTTTATGGAATAAATG 207530
QY 716 lulysValGlyValProThrIleGlnGlnLeuLeuGluTrpSerPheIleAsnSerAsnL 736
Db 207531 AGAAAGTCGGAGTACCACAAATCCACAGATTATTAGAAATGTTCTTTTATCAACAGCAACC 207590
QY 736 eulysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspP 756
Db 207591 TGAATTTTCGGAGGACCATCATGTTGTATATCCAGATGCCCTTGGTTTGGGAAGACT 207650
QY 756 helysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluA 776
Db 207651 TTAACACTATTAAAAAATTTTTCCTCCCTCGAATTAATATACAGATTTTACTTTGAAG 207710
QY 776 spThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysT 796
Db 207711 ACACCTCCAGGAGTCCGACATCTGTGGAGGACTCGCCATGTATGAGTGTAGAGAGTGCT 207770
QY 796 yrAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrG 816
Db 207771 ATGATGCCCGGACATCTCGCAGGGAAGATCAAGCAGTTCTGTAAAGACCTGCAGCACTC 207830
QY 816 lnValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysA 836
Db 207831 AGGTTACACTTCATCCAGAGACTGAATCACACTTACCATCCAGTATCACATCCCAAG 207890
QY 836 spLeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheA 856
Db 207891 ACTTGCCCGACTGGGACTGGAGACATGGCTGCATCCCATGTCAAGAGATGGAGTTATTG 207950
QY 856 lavalLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspS 876
Db 207951 CTGTGCTCTGCATAGAAACCCAGCCACTATGTTGCTTTTGTGAAGTACGGGAAGGATGACT 208010
QY 876 erAlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnI 896
Db 208011 CTGCCTGGCTCTCTTTTGACAGCATGGCTGATCGAGATGGTGTGATGAGATGGCTTCAACA 208070
QY 896 leProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspL 916
Db 208071 TTCCACAAGTGACACCCCTGCCCAAGAGTAGGAGAGTACTTTGAAGATGTCTCTGGAGGACC 208130
QY 916 euHisSerLeuAspSerArgArgIleGlnGlyCysAlaArgLeuLeuCysAspAlaT 936
Db 208131 TGCACCTCTTTGGACTCCAGAAGGATTCAAGGCTGTGGCGCAGACTTCTTTGGGATGCAT 208190
QY 936 yrMetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
Db 208191 ACATGTGCATGTACACAGATCCAAACCATGAGCTTGTACAAA 208231
```

Search completed: April 18, 2005, 18:28:45
Job time : 9311 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 15:23:45 ; Search time 54 Seconds

(without alignments)
1242.248 Million cell updates/sec

Title: US-09-671-687A-3

Perfect score: 5034

Sequence: 1 MSSGLWSQEKVTSFYWEERI.....RLLCDAYMCVQSPWLSLYK 949

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 308579 seqs, 70686408 residues

Total number of hits satisfying chosen parameters: 308579

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pcp.*

2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pcp.*

3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pcp.*

4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pcp.*

5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pcp.*

6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pcp.*

7: /cgn2_6/ptodata/1/paa/US11_NEW_COMB.pcp.*

8: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4971.5	98.8	956	US-10-760-678-4	Sequence 4, Appli
2	122.5	2.4	1499	1 PCT-US04-24868-15	Sequence 15, Appl
3	122	2.4	2296	6 US-10-696-909A-46	Sequence 46, Appl
4	122	2.4	2752	6 US-10-696-909A-44	Sequence 44, Appl
5	120.5	2.4	915	8 US-60-655-875-141720	Sequence 141720,
6	119.5	2.4	1451	1 PCT-US05-03344-1	Sequence 1, Appli
7	119.5	2.4	1451	7 US-11-046-346-1	Sequence 1, Appli
8	119.5	2.4	2412	8 US-60-643-717-7896	Sequence 7896, Ap
9	115	2.3	2543	6 US-10-450-763-51654	Sequence 51654, A
10	113	2.2	1181	8 US-60-664-582-558	Sequence 558, App
11	113	2.2	1195	8 US-60-659-397-1215	Sequence 1215, Ap
12	113	2.2	1200	8 US-60-664-582-563	Sequence 563, App
13	113	2.2	1213	7 US-11-033-545-572	Sequence 572, App
14	113	2.2	1213	6 US-60-659-397-1214	Sequence 1214, Ap
15	113	2.2	1226	7 US-11-033-545-313	Sequence 313, App
16	113	2.2	1226	8 US-60-664-582-562	Sequence 562, App
17	113	2.2	1226	8 US-60-664-582-564	Sequence 564, App
18	112.5	2.2	1116	6 US-10-818-066-31	Sequence 31, Appl
19	112	2.2	807	6 US-10-450-763-45311	Sequence 45311, A
20	112	2.2	869	7 US-11-031-175-10125	Sequence 10125, A
21	112	2.2	3684	1 PCT-US05-01768-8	Sequence 8, Appli
22	111.5	2.2	1205	6 US-10-489-448-1764	Sequence 1764, Ap
23	111.5	2.2	1304	8 US-60-651-509-263	Sequence 263, App
24	111.5	2.2	1304	8 US-60-651-235-1508	Sequence 1508, Ap
25	111.5	2.2	1304	8 US-60-664-579-738	Sequence 738, App

26	111	2.2	1673	8 US-60-643-717-14236	Sequence 14236, A
27	110	2.2	1497	6 US-10-450-763-51156	Sequence 51156, A
28	110	2.2	2644	6 US-10-450-763-50462	Sequence 50462, A
29	109.5	2.2	1304	8 US-60-651-509-262	Sequence 262, App
30	109.5	2.2	1304	8 US-60-651-235-1507	Sequence 1507, Ap
31	109.5	2.2	1304	8 US-60-664-579-737	Sequence 737, App
32	109	2.2	452	1 PCT-US04-43356-148	Sequence 148, App
33	109	2.2	452	7 US-11-021-949-148	Sequence 148, App
34	109	2.2	452	7 US-11-021-928A-45	Sequence 45, Appli
35	109	2.2	521	7 US-11-044-879-4	Sequence 4, Appli
36	109	2.2	866	1 PCT-US04-17965-1594	Sequence 1594, Ap
37	109	2.2	866	1 PCT-US04-17965B-1594	Sequence 1594, Ap
38	109	2.2	866	1 PCT-US04-17965C-1594	Sequence 1594, Ap
39	109	2.2	866	6 US-10-863-905-1594	Sequence 1594, Ap
40	108.5	2.2	493	8 US-60-651-509-442	Sequence 442, App
41	108.5	2.2	493	8 US-60-651-235-1229	Sequence 1229, Ap
42	108.5	2.2	703	7 US-11-096-039-11	Sequence 11, Appl
43	108.5	2.2	809	6 US-10-450-763-40966	Sequence 40966, A
44	108.5	2.2	1205	6 US-10-450-763-52696	Sequence 52696, A
45	108	2.1	714	6 US-10-450-763-45506	Sequence 45506, A

ALIGNMENTS

RESULT 1

US-10-760-678-4

; Sequence 4, Application US/10760678

; GENERAL INFORMATION:

; APPLICANT: Derr, Jonathan

; APPLICANT: Fanslow, William

; APPLICANT: Dougall, William

; TITLE OF INVENTION: SCREENING ASSAYS FOR AGONISTS OR ANTAGONISTS OF CD40 SIGNALING

; FILE REFERENCE: 3198

; CURRENT APPLICATION NUMBER: US/10/760, 678

; CURRENT FILING DATE: 2004-01-20

; PRIOR APPLICATION NUMBER: US/09/851, 673

; PRIOR FILING DATE: 2001-05-08

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: Patent in version 3.0

; SEQ ID NO 4

; LENGTH: 956

; TYPE: PRT

; ORGANISM: Homo sapiens

; US-10-760-678-4

Query Match	98.8%	Score 4971.5	DB 6	Length 956
Best Local Similarity	99.2%	Pred. No. 0		
Matches	948	Conservative	0	Mismatches
			1	Indels
			7	Gaps
QY	1	MSSGLWSQEKVTSFYWEERIF	YLLQLQEC	SVTDKQTKLLKVPKGSIGQYIQDRSVGHSRI
				60
Db	1	MSSGLWSQEKVTSFYWEERIF	YLLQLQEC	SVTDKQTKLLKVPKGSIGQYIQDRSVGHSRI
				60
QY	61	PSAGKKNQIGLKILQPHAVLPVDE	DDVINEKFTTELLA	ITNCBERSLFFKNRRLS
				119
Db	61	PSAGKKNQIGLKILQPHAVLPVDE	DDVINEKFTTELLA	ITNCBERSLFFKNRRLS
				120
QY	120	KGLQIDVGPVKVQLRSGBEKG	PGVVRFRGPLAERTVSGI	FFGVVLELLEGRGGQFTDGV
				179
Db	121	KGLQIDVGPVKVQLRSGBEKG	PGVVRFRGPLAERTVSGI	FFGVVLELLEGRGGQFTDGV
				180
QY	180	YQKQLFQDCEDCG	FVALD	KLELIEDDDTALSDYAGPGDTM
				238
Db	181	YQKQLFQDCEDCG	FVALD	KLELIEDDDTALSDYAGPGDTM
				240
QY	239	GETIESGTWIFCDVLPKESLGY	FVGVDMNDP	IGNWDGREDGV-LCSFACVESTILLHN
				297
Db	241	GETIESGTWIFCDVLPKESLGY	FVGVDMNDP	IGNWDGREDGV-LCSFACVESTILLHN
				300
QY	298	DIIP----	ESVTQRRRP	PKLAFMRSRGVDKGSSHNKPKATGSTD
				353
Db	301	DIIPALSES	VTQRRRP	PKLAFMRSRGVDKGSSHNKPKATGSTD
				360

QY 302 ESWTQERRPPKLAFFMSRGVGDGSSS-----HNPKATGTSF-----PGNRSELFTYTLNG 353
Db 1163 ESKEMALPPQEDATAPPRQDKFPFPVQDRPSSSLVPKDTLRTPPRRSGA-----G 1217
QY 354 SSVDSOPQ-----SKSKNTWYIDEVAEDPA-----KSLTEISTDFDRSSPPLQPP 398
Db 1218 SSPETKEQNSALPTSQDELMVEVKSSEPAQILSHLSSELKEMSTNSFESSPEVER 1277
QY 399 PVNSLTENRHFSLPFLTKMPNTNGSIGHSPLSLSAQSVMEBLNAPVQES---PPLAM 455
Db 1278 PAVSLTLDQSQASLEAVEVPSMASSWGGPHFS-----PEHKELNSPLRENSFGSPLEF 1333
QY 456 PPGNSHGLEV--GSLAEVKE--NPPF 477
Db 1334 RNSGPLGTEMNTGFSSEVKEDLNGPF 1359

RESULT 4

US-10-696-909A-44
; Sequence 44, Application US/10696909A
; GENERAL INFORMATION:
; APPLICANT: Lorens, James B.
; APPLICANT: Atchison, Robert E.
; APPLICANT: Frieria, Anabella
; APPLICANT: Holland, Sacha
; APPLICANT: Rigel Pharmaceuticals, Inc.
; TITLE OF INVENTION: Modulators of Angiogenesis and Tumorigenesis
; FILE REFERENCE: 021044-005820US
; CURRENT APPLICATION NUMBER: US/10/696,909A
; CURRENT FILING DATE: 2003-10-29
; PRIOR APPLICATION NUMBER: US 60/512,251
; PRIOR FILING DATE: 2003-10-17
; PRIOR APPLICATION NUMBER: US 60/421,989
; PRIOR FILING DATE: 2002-10-29
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 44
; LENGTH: 2752
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: splicing coactivator subunit SRM300; RNA binding
; OTHER INFORMATION: protein; AT-rich element binding factor
US-10-696-909A-44

Query Match 2.4%; Score 122; DB 6; Length 2752;
Best Local Similarity 22.8%; Pred. No. 7.8;
Matches 88; Conservative 53; Mismatches 167; Indels 78; Gaps 16;
QY 128 CPVKVQLRGEEKFPQGVWFRGPLLAERTVSGIFFGVVLELLE-RGRGQGFTDGVYQKQLF 186
Db 1016 CQEQSKSLVQSCPSGLSACGVKSTPGEYFVGVSLLQKQSQTSQSPDH----- 1067
QY 187 QCDEDCGFVALDKLEIIDDATTALESDDYAGPGDTMQ--VELPPLINSRVSLKGETIES 244
Db 1068 -----RSDTSSP-EVRQSHSESPSLQSKQTSQSPKGRSRSS 1102
QY 245 GTVI-PCDVLPGKESIGYFVGVDN-DNPIGNWDRPDGVLCSFACVESTILLHINDI-IP 301
Db 1103 SPVTELASRPIQRDRGEFSASPLKSGMSPEQSRFQSDSSSYPTVDSNLLQSQSRLETA 1162
QY 302 ESWTQERRPPKLAFFMSRGVGDGSSS-----HNPKATGTSF-----PGNRSELFTYTLNG 353
Db 1163 ESKEMALPPQEDATAPPRQDKFPFPVQDRPSSSLVPKDTLRTPPRRSGA-----G 1217
QY 354 SSVDSOPQ-----SKSKNTWYIDEVAEDPA-----KSLTEISTDFDRSSPPLQPP 398
Db 1218 SSPETKEQNSALPTSQDELMVEVKSSEPAQILSHLSSELKEMSTNSFESSPEVER 1277
QY 399 PVNSLTENRHFSLPFLTKMPNTNGSIGHSPLSLSAQSVMEBLNAPVQES---PPLAM 455
Db 1278 PAVSLTLDQSQASLEAVEVPSMASSWGGPHFS-----PEHKELNSPLRENSFGSPLEF 1333

QY 456 PPGNSHGLEV--GSLAEVKE--NPPF 477
Db 1334 RNSGPLGTEMNTGFSSEVKEDLNGPF 1359

RESULT 5

US-60-655-875-141720
; Sequence 141720, Application US/60655875
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey
; APPLICANT: Du, Zijing
; APPLICANT: Guo, Liang
; APPLICANT: Kovalic, David
; APPLICANT: Lu, Maolong
; APPLICANT: McCarter, James
; APPLICANT: Miller, Nancy
; APPLICANT: Williams, Deryck
; APPLICANT: Vaudin, Mark
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETERODERA INFESTATIONS
; TITLE OF INVENTION: IN PLANTS AND COMPOSITIONS THEREOF
; FILE REFERENCE: 38-21(53885)
; CURRENT APPLICATION NUMBER: US/60/655,875
; CURRENT FILING DATE: 2005-02-24
; NUMBER OF SEQ ID NOS: 171306
; SEQ ID NO 141720
; LENGTH: 915
; TYPE: PRT
; ORGANISM: Heterodera glycines
; FEATURE:
; OTHER INFORMATION: Coding regions on vcdna: vcdna-seqID 68143; Strand=-; Position=1
; OTHER INFORMATION: -348,394-905,955-1966,2010-2170,2218-2308,2353-2976
; FEATURE:
; OTHER INFORMATION: Homolog annotation: Hit ID=NP_001004509.1; Match level="QueryCover
; OTHER INFORMATION: 87%; HitCoverage=55%; E-value=-1e-111; Identity=32%; Hit description
; OTHER INFORMATION: =si:bouseml-18005.3 [Danio rerio] emb[CAD60788.1] novel protein sim
; OTHER INFORMATION: to DNA polymerases [Danio rerio]
US-60-655-875-141720

Query Match 2.4%; Score 120.5; DB 8; Length 915;
Best Local Similarity 21.9%; Pred. No. 2;
Matches 106; Conservative 69; Mismatches 195; Indels 115; Gaps 23;
QY 498 ELEDECAGC--TDGTFRGTRYFTCALKKALFKVCLKSCRPDSRFASLQPVNSQIE----- 549
Db 49 ELSTDCEICCTQNGTFERTKEWLSKLSKFIFYQNSNRSEAD-GDEQPVAAFVWILRAW 107
QY 550 -----RCNSLAFQGY--LSEVVEENTPPKMEKEGLEIMI GKKGIGQHYNSCYLDS 598
Db 108 NNQYKTYIAHNASRFDGHEALNYICKTVRRPDVVMNGLKIYEFVRHSPKSHMLIMRDS 167
QY 599 TLFCLFAFSSVLTLLRPKEKNDVEYSETQE--LLRTEIVNPLRIY--GYVCATKIMK 654
Db 168 CLIMPIPLADLKTTFNLDCEDKPFPPYAFNRRENYCIRLEKLPDQKMYEPPGSMKADKYEK 227
QY 655 LRKILEKVEAASGFTSEE-----KDPEEFNLIL--FHHIL-----RVEPL-LKIRS 697
Db 228 FKWYGENEDTPPLPEQLRTYCHNDTEILLKALVFRHILINKITQGFVLPISCTTAS 287
QY 698 AGQKQDCVFPYQ-----IFMEKNEKVGVPITQOLLEWSFINSLKFAEAPSLIIQ 748
Db 288 ACNNIFKAQPMQEDQLAMVPELGYERNDRAVSIAI-KYLDWRAKSEGIDIQHAGN----- 341
QY 749 MPFRG--KDPKLPKFKIPPSLEMLNITLLEDTPQCRCICGGLAMYTECREDYDDDISAKI 806
Db 342 ----GREKQWKFEK-----LDGWIESQQRCEIEVIGCYWHGCDRCF-KPD----- 380
QY 807 KQFC--KTCNTQVHLHPKRLNHNKYNPVSPLPKDLPDWDWRHGCCIPCQNMELFAVLCIETSH 864
Db 381 EQLVDEKTCOELNEMTQDRLRLQRE-----PDTDGR-----SLQVBEIWECEIIND 425
QY 865 YVAFVKYKDDSAWLFPDMSMADRGQGNFIPQVTPC-----PEVGEYLKMSLE 914

Db 426 QLK-----KNHEMKVFFDLDLANERGLD-----PRLAYCGGRTGTLRLFAEASADEKISVF 476
Qy 915 DLHSL 919
Db 477 DIVSL 481

RESULT 6
PCT-US05-03344-1
; Sequence 1, Application PC/TUS0503344
; GENERAL INFORMATION:
; APPLICANT: Dana Farber Cancer Institute
; APPLICANT: D'Andrea, Alan D
; TITLE OF INVENTION: Method for Determination and Quantification of Radiation or Genot
; TITLE OF INVENTION: Exposure
; FILE REFERENCE: 7032/2082
; CURRENT APPLICATION NUMBER: PCT/US05/03344
; CURRENT FILING DATE: 2005-01-28
; PRIOR APPLICATION NUMBER: US 60/540380
; PRIOR FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1451
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-03344-1

Query Match 2.4%; Score 119.5; DB 1; Length 1451;
Best Local Similarity 18.3%; Pred. No. 4.7;
Matches 198; Conservative 141; Mismatches 376; Indels 367; Gaps 47;

Qy 3 SGLWSEKVTSPYWEERIFVLLQOE-CSVTDKQTKLLK-----VPGKSI-----46
Db 409 SCGIOEQLLQSTF---SVHYLVLMKCMSSILSLAQSLHSLDQSIISFGSLLYKYAFKFF 465
Qy 47 GOYIQRDSVGHSHRIPSAKGNKQI--GLKILEQPHAVLFVDEVDWEINEKFTLELL-----100
Db 466 DTYCQOEVEGVALVTHICSGNEAEVDALDVLE---LVVLNPSAMMNAVFQGIIDYLD 522
Qy 101 --AITNCEERFSL-----FKNRNLSKGLQIDVGCVPKVLRSGBEKKFP-----GVVRF 148
Db 523 NISPOQIRKLFYVLSLAFSKQNEASSHIQDDMHLVIRKQLSSTVPKYKLGIGIAVTMA 582
Qy 149 GPILLARTVSGIFFGVELLEBRGQGTGVQKOLFQ-----C-----DEDC 192
Db 583 GIMAADRSSES-----PSLTQERANLSDBQCTQVTSLLQLVHSCRSRQSPQASALYYDEFA 636
Qy 193 GFVALDKLE-----LIEDDDTALESYDAGP-----GDTWQ 222
Db 637 NLIQHEKLPKALEWGHGTCNDQFQAVVDSCVPEGDPPPVPKALYGLBEYDTQDGA 696
Qy 223 VELPPLPINSRSLKGGETI--ESG-----SLGYF-----245
Db 697 INLLPFLFSQDPAKDGPTVTSQESGKLYSPICLAPYFRLLRLCVERQHNGNLEEIDGLL 756
Qy 246 --TVIFCDVLPKBE-----262
Db 757 DCPIFLTDLEPGKLESMGAKESFMCSLIFLTLNFWREIVNAFCQETSPENKGVLTREL 816
Qy 263 -----VGVDMDNPIGNWGDGFDGVLCSFACVESTILLHINDIIPESVT--305
Db 817 KHIVELQILLEKYLAVTPDPVPLGNFD-----VETL-----DITPHVTAI 858
Qy 306 --QERPPKALFWSRGVGDGKSSHNKPKATGTSBP-GNRRSELFTYTLNGSSVDSPQS 362
Db 859 SAKIRKKGIERKQKTGSKTSSDLSLSEKNSECPTPSHRQLKNEFTGKEKTSLLL 918
Qy 363 KSKNTWYID-----EVAEDPAKSLTISTDFDRSSPP 394
Db 919 HNSHAFPRELDIEVFSILHCGLVTKFLTDPMTEATEVVLQGPPELFLLEDLSOKLES 978

Qy 395 LQPPVNSLTTENRPHSLPFSLTWKMPNTNGSIGHSPL-SLSAOSV---MEELNTAPVQES 450
Db 979 MLTPPI-----ARRVPEFLKNK---GSRNIGFSHLQORSQAQEIHVHVEQLLT-----1021
Qy 451 PPLAMPNGN-----SHGLEVSGSLAEVKEN-----PPFYGVIRWTIG--Q 486
Db 1022 -PMCNHLENIHNYIOCLAAENHGVVDGPGVKVQOEYHIMSSCYORLLQIFHGLFAMSGFSQ 1080
Qy 487 PPGLNEVLAGL-----ELEDECAGCTDGTFRGTRYFTCALKKALFVKL 529
Db 1081 PENQNLLYSALHVLSSRLKQSGHSQPLEBLLSQSVHYLQNFHQSPISFOCAL---YLIRL 1137
Qy 530 KSCRPDSRFASLOPVSNQIERCNSLAFGGYLSEVVEENTPPKMEKEGLIMICKKGIQ 589
Db 1138 LMVILEKTASAQ-----NKEKIASLA-RQFLCRVM-----PSGDKESNI-----SNDQL 1182
Qy 590 HYNSC-YLDSTLFLCLFAPSSV---LDTVLLRPKKNQDVVEYSETOELLRTTEIVNPLRIY 644
Db 1183 HALLCIYLEHTESILKAEIEIAQVGPVELINPKDASSSTFPLTRH-----TPVVF 1235
Qy 645 GYVCATKIMKRLKILEKVEAASGFTSEKDPBEFL--NILFHILRVEPLLKIRSAGKV 702
Db 1236 RVWMA-----ELEKIVKIEPGTAADSQQIHEEKLLYWNMAVRDFSILINLIKVDGHPVL 1291
Qy 703 QDCY-FYQIFMEKNEKVGUPTIQ-----OLLEWSFINSNLKFAEAPSCILLIOMP 750
Db 1292 HVCLKVGRLFVEAFKLQCMPLLDISFRKHREDVLSLLETFTOLDTRLHLHLCGHSKHODT 1351
Qy 751 REGKDFKLFKIFPSPLELNITDLEDTTPQCRCICGLAMVECRECY-----DDPDISAGK 805
Db 1352 RUTQHVPLKK--TLELLV-----CRVKAMLTUNNCREAFWLGNLKNRDLQGE 1398
Qy 806 IK 807
Db 1399 IK 1400

RESULT 7
US-11-046-346-1
; Sequence 1, Application US/11046346
; GENERAL INFORMATION:
; APPLICANT: Dana Farber Cancer Institute
; APPLICANT: D'Andrea, Alan D
; TITLE OF INVENTION: Method for Determination and Quantification of Radiation or Genot
; TITLE OF INVENTION: Exposure
; FILE REFERENCE: 7032/2082
; CURRENT APPLICATION NUMBER: US/11/046,346
; CURRENT FILING DATE: 2005-01-28
; PRIOR APPLICATION NUMBER: US 60/540380
; PRIOR FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1451
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-046-346-1

Query Match 2.4%; Score 119.5; DB 7; Length 1451;
Best Local Similarity 18.3%; Pred. No. 4.7;
Matches 198; Conservative 141; Mismatches 376; Indels 367; Gaps 47;

Qy 3 SGLWSEKVTSPYWEERIFVLLQOE-CSVTDKQTKLLK-----VPGKSI-----46
Db 409 SCGIOEQLLQSTF---SVHYLVLMKCMSSILSLAQSLHSLDQSIISFGSLLYKYAFKFF 465
Qy 47 GOYIQRDSVGHSHRIPSAKGNKQI--GLKILEQPHAVLFVDEVDWEINEKFTLELL-----100
Db 466 DTYCQOEVEGVALVTHICSGNEAEVDALDVLE---LVVLNPSAMMNAVFQGIIDYLD 522
Qy 101 --AITNCEERFSL-----FKNRNLSKGLQIDVGCVPKVLRSGBEKKFP-----GVVRF 148
Db 523 NISPOQIRKLFYVLSLAFSKQNEASSHIQDDMHLVIRKQLSSTVPKYKLGIGIAVTMA 582
Qy 149 GPILLARTVSGIFFGVELLEBRGQGTGVQKOLFQ-----C-----DEDC 192
Db 583 GIMAADRSSES-----PSLTQERANLSDBQCTQVTSLLQLVHSCRSRQSPQASALYYDEFA 636
Qy 193 GFVALDKLE-----LIEDDDTALESYDAGP-----GDTWQ 222
Db 637 NLIQHEKLPKALEWGHGTCNDQFQAVVDSCVPEGDPPPVPKALYGLBEYDTQDGA 696
Qy 223 VELPPLPINSRSLKGGETI--ESG-----SLGYF-----245
Db 697 INLLPFLFSQDPAKDGPTVTSQESGKLYSPICLAPYFRLLRLCVERQHNGNLEEIDGLL 756
Qy 246 --TVIFCDVLPKBE-----262
Db 757 DCPIFLTDLEPGKLESMGAKESFMCSLIFLTLNFWREIVNAFCQETSPENKGVLTREL 816
Qy 263 -----VGVDMDNPIGNWGDGFDGVLCSFACVESTILLHINDIIPESVT--305
Db 817 KHIVELQILLEKYLAVTPDPVPLGNFD-----VETL-----DITPHVTAI 858
Qy 306 --QERPPKALFWSRGVGDGKSSHNKPKATGTSBP-GNRRSELFTYTLNGSSVDSPQS 362
Db 859 SAKIRKKGIERKQKTGSKTSSDLSLSEKNSECPTPSHRQLKNEFTGKEKTSLLL 918
Qy 363 KSKNTWYID-----EVAEDPAKSLTISTDFDRSSPP 394
Db 919 HNSHAFPRELDIEVFSILHCGLVTKFLTDPMTEATEVVLQGPPELFLLEDLSOKLES 978

Qy	149	GPLAERTVSGIFPGVELLEBEGRGQFTDGVGYQKQLFQ	-----C-----	DEDC	192	
Db	583	GIMAARDS-	-----PSLTQERANLSDEQCTQVTSLLQLVHSCSEQSPASALYDEFA	636		
Qy	193	GFVALDKLE	-----LIEDDDTALESYAGP	-----	GDTMQ 222	
Db	637	NLTQHEKLDPKALEWVGHTICNDPQDAFVVDVSCVVPPEGDFPFVKALYGLEEYDTQDGIA	-----	696		
Qy	223	VELPPLEINRSVLKGETI	--ESG	-----	245	
Db	697	INLLPLLFSDFAKDGGPVTSQESGGKLVPLCLAPYFRLLRLCVRFOHNGNLEEIDGLL	-----	756		
Qy	246	--TVIFCDVLPGKE	-----SLGYF	-----	262	
Db	757	DCPIFTLDLEPGKLESMASAKFMSGLIFTLINWFREIVNAPCQETSPENKGVLTRL	-----	816		
Qy	263	-----	-----	-----	305	
Db	817	KHIVELOILEKYLAVTPDVPPLGNFD	-----VETL	-----	DITPHTVTAI 858	
Qy	306	--QERRPPKLAFRSGVDGSGSSHNNPKATGSTDP	-----GNRRSELFYTLNGSSVDSQPQS	362		
Db	859	SAKIRKKGKIERKQKTGKTSSTSSDTLSEBKSECPTPSHRGQLNKKEFTGKEKTSLLL	-----	918		
Qy	363	KSXNTWID	-----EVAEDPAKSUTEISTDPRSSPP	394		
Db	919	HNSHAPFREILDIEVFSILHCGLVTKFILDTEHTEATEVVQLGPPELLFLELDLSQKLES	-----	978		
Qy	395	LQPPVNSLTENRPHSLPSLTMPNNGSIGHSP	-----SLSAQSV	-----MEELNAPQVES	450	
Db	979	MLTPPI	-----ARRVPLKNN	-----GSRNIGFSLHQRSQAIEVHCVEQLLIT	1021	
Qy	451	PPLAMPNGN	-----SHGLEVGSLAEVKEN	-----PPFYGVIRWIG	-Q 486	
Db	1022	-PNCNHLNTHNYIOCLAEHNHGVDPGVKVQOEYHIMSSCYQRLLOIFHGLFANSGFSQ	-----	1080		
Qy	487	PPGLNEVLACL	-----ELEDECAGCTGCTFRGTRYFTCALKALFVKL	529		
Db	1081	PENQNLILYALHVLSSRLKQGEHSQPLEELLSQSVHYLQNFHQSI	-----PSFOCAL	-----YLIRL	1137	
Qy	530	KSCRPDSRFASLPVNSQIERCNSLAFFGYLSEVVENTPPEKEGLEITWICKKGIQ	589			
Db	1138	LMVILEKSTASAO	-----NKEIASLA	-----ROFLCRW	-----PSGDKEKSN	-----SNDQL 1182
Qy	590	HYNSC	-----YLDSTLFCLPFASV	-----LDTVLLRPEKXNDVEYSETOELLRTETVNPRLRY	644	
Db	1183	HALLCILEHTESILKAIEBIAQVGPPELINSKDASSSTFPLTRH	-----	-----TPVVF	1235	
Qy	645	GYVCATKIMKRIKLEKVEAASGFTBEEKDPBEFL	-----NILFHHILRVEPLLKIRSAQKV	702		
Db	1236	RVMWA	-----ELEXIVKIEBPGTAADSQOIHEEKLLYNMAVRDFSILINLIKIVDFDHPVL	1291		
Qy	703	QDCY	-----FYQLFMENEKVGP	-----QLEWFSFINSNLKFAEAPSLIOMP	750	
Db	1292	HVCLKVGRFLVEAFLLKQCMPLDISFRKHREDVLSLET	-----POLDRLLHLCGHSKHQDT	1351		
Qy	751	RFKDFKLFKKIPPSLELNITDLETPROCRI	-----CGGLAMVECRECY	-----DDPDISAGK	805	
Db	1352	RLTQHVPLLK	-----TLELLV	-----CRVKAMLTNNCREAFWLGNLKNRDLQGE	1398	
Qy	806	IK	807			
Db	1399	IK	1400			

```

RESULT 8
US-60-643-717-7896
; Sequence 7896, Application US/60643717
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S.
; TITLE OF INVENTION: Genes and Uses for
; FILE REFERENCE: 38-21(53629)A

```

```

; CURRENT APPLICATION NUMBER: US-60/643,717
; CURRENT FILING DATE: 2005-01-12
; NUMBER OF SEQ ID NOS: 19247
; SEQ ID NO 7896
; LENGTH: 2412
; TYPE: prt
; ORGANISM: ASPERGILLUS NIDULANS FGSC A4
US-60-643-717-7896

```

Query Match	2.4%;	Score 119.5;	DB 8;	Length 2412;
Best local similarity	19.7%;	Pred. No. 9.9;		
Matches 113;	Conservative 81;	Mismatches 212;	Indels 169;	Gaps 28;
QY	330	KPKATGTS	DPGNRRSE	-----LFYTLNGS-----SVDSPQSKSKNTWY-----370
DB	292	EPAREDT	PDGPGVAL	PEGALATLOFTPLEGTEPWSSETSP-----NEEMQOVAAPPERRL348
QY	371	DEVAEDP	AKSLTEIS	-----TDFDRSSPPLQPPPPVNSLTTERRFHS411
DB	349	DD--ETP	VELSEY	FTGLSIPSPGGFFNSLAPRARHTWSLPKLNQPPPTSA--TAERFFYN404
QY	412	LPESLT	KMPNTNGS	IGHSPLSLSAQSVMEELNAPVQESPPPLA--MPPGNSHGLEVGS LA469
DB	405	LPFN-----	REGEIIE	VIDUPERLNDQLTAI-----YAPPTAIKIPESPAHPPTEGSIS456
QY	470	EYKENP	PPFYGVIR	WIQOP-----PGLNEVLAGLEDEDECAGCTYDGT---FRGTRYETFCAL521
DB	457	PVSEY-----	VHEISRP	TAYVDPDQDENYAEELHKKALSSLDRTSVMLAAQASYLAAL510
QY	522	KKALFVKL---	KSCR	PDSPFASLPQVSNQIYERCNLSIAFGYILSEVVVEENTPPKMEKE---575
DB	511	RETNPV	NNLPDEDER	PDQVDESPQHVSPALERNASVCFTCMPFPPSPSLPAANASKDSIY570
QY	576	--GLE	TIMCKK-	-----GIGHVNSCYLDSTLFCI-----F604
DB	571	WGFREF	LQGSRRDT	FVHRSTRFDVAQSFRLGSLGHKKLLGNLTVELVLPDRPAYSGPF630
QY	605	AFS-----	SVLDTV	LLRPKKNDEYVYXSETQELLRTETVNP---LRIYGYVCATKIM---KL655
DB	631	AKAPRHS	VLPGIL	LQKAEFSMI-----EKEQLVLSQISQPMWAMEALRYLQGGNLVVS PAR686
QY	656	RKILEK	VEAASGFT	SEEKDPBEFLNLFH-----HILRVPELLKIRSAQGVQDCYFY708
DB	687	KGFSKR	ATAAP	PHKTPKREQVRVLDLGGHATAEAWHMLAHDPYHVKV-----Y734
QY	709	QIFMEK	NEKVGVP	TIOQLLEWSFINSNLKFAEAPSCLIIQMRFPGDKFLFKKIFPSLEL768
DB	735	TVYTEHQ	-----	VNKAIK--GPPNHRHIQVP-----QLWKLPFFDNKF771
QY	769	NITD-----	LLEDTP	-QCRICGGLAWYECREC795
DB	772	DVISAR	SLPAFLK	TERPAGDCLDEYDCLCKERCRC806

```

RESULT 9
US-10-450-763-51654
, Sequence 51654, Application US/10450763
, GENERAL INFORMATION:
, APPLICANT: Hyseq, Inc
, TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
, FILE REFERENCE: 790CIP3/US
, CURRENT APPLICATION NUMBER: US/10/450,763
, CURRENT FILING DATE: 2003-06-11
, PRIOR APPLICATION NUMBER: PCT/US01/08631
, PRIOR FILING DATE: 2001-03-30
, PRIOR APPLICATION NUMBER: 09/540,217
, PRIOR FILING DATE: 2000-03-31
, PRIOR APPLICATION NUMBER: 09/649,167
, PRIOR FILING DATE: 2000-08-23
, NUMBER OF SEQ ID NOS: 60736
, SOFTWARE: Custom
, SEQ ID NO 51654
, LENGTH: 2543

```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1093)..(1110)
; OTHER INFORMATION: PWWP domain proteins domain identified by eMATRIX, accession
; OTHER INFORMATION: number PF00855, p-value=5.605e-18, raw score of 13.75
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1088)..(1188)
; OTHER INFORMATION: PWWP domain identified by PFam, accession name PWWP, E-value=
; OTHER INFORMATION: 2.6e-38, PFam score of 140.7
US-10-450-763-51654

```

Query Match	2.3%;	Score 115;	DB 6;	Length 2543;	
Best Local Similarity	18.8%;	Pred. No. 23;			
Matches 120;	Conservative 72;	Mismatches 193;	Indels 254;	Gaps 26;	
Qy	190	EDC-----GVALDKLELI-----	EDDDTALESYAGPGTQMVLPPLNSRV	234	
Db	1515	KDCPLQFHDFKSVDLKVCPRYTAVLASEDDGIGIE-----	LDTLQLETLTSSASR	1569	
Qy	235	SLKGGETTESGTVIFCDVLPKGESLYGVGVDMDNPIGNWDRFGDVLCSFACVESTILL	294		
Db	1570	RLR---VLEAETQILTD-WQDKGDRREKLGRDHGLG-----	1603		
Qy	295	HINDIIPESVTQERRPPKPLAFMRGVDGKSSSHNPK-----	ATGSTDPGNRRSELF	348	
Db	1604	-----APPK-----	HCKPKKQLEKGAGHGGPGGR---	1630	
Qy	349	YTLNGSSVDSQFSKSKNTWIYDEVAEDPAKSLTEISTDFRSPPLPPPPVNSUTTENR	408		
Db	1631	-----PKSKN-----	LQPKIOEYETDDPIDVPRIKNDAPNR	1663	
Qy	409	FHSLPFSLTUKMPTNGSICHSPLSLSAQSV---MEELINTAPVQESPLAMPNGSH-----	461		
Db	1664	FWA-----	SVPEYCADITSEBVRTLEELKPPDEAEHYKIPPLGKHYSQRW	1710	
Qy	462	-----GLEVGSLAEVKENPPFYGVIRWIGQPPGLNEVLAGLELEDECAGCTDGT	510		
Db	1711	AOEDLLEQDKGARAANVADKK-----	GLMGPLFELDTKDVA-----	1749	
Qy	511	FRGTRYFTCAKKALFVKLKSRCPSRFSASLPVSNQIERCNSLAPGG-----	YLSEVVEE	566	
Db	1750	-----LLKK-----	SEAQHEQPE-----	DGCPFGALTQLLOALVEE	1781
Qy	567	NTPPMWEKEGLEIMIGKKKGIOGHVNSCYLDDSTLFCLEFASFSSVLTIVLLRPKEKD-----	622		
Db	1782	NIISPMDSPIDMSGKESGADGASTSPRNQKNPF-----	SVPTKSLSEIKELIAQ	1835	
Qy	623	--VEYVSETQELLRTIEIVNPLRIYGVVCATKIMLKRLKILEKVEAASGFFSEBKDPEERLN	680		
Db	1836	GLLESEDRPAEDSEDEV-----	LAELRKQAEULKAHNRRTKK--HDLUR	1879	
Qy	681	ILFHILRIVEPILKIRSAQKQVQDCYFYQIFMEKNEKVGUPT-----	722		
Db	1880	LAKEVSRQELRQRYMADNEWMDA--	PFKINAAEQKKETPTKKEKQAWKTKSERSTL	1937	
Qy	723	-----IQOLLEWSFNINSLKFAEAPSLIIQMPRFGKDF	756		
Db	1938	KLLDGVNALOITFNTNRVERIE-RCLEL-----	FGRDY	1971	

```

, FILE REFERENCE: LC001470
, CURRENT APPLICATION NUMBER: US/60/659,397
, CURRENT FILING DATE: 2005-03-09
, NUMBER OF SEQ ID NOS: 47859
, SOFTWARE: FastSEQ for Windows Version 4.0
, SEQ ID NO 1215
, LENGTH: 1195
, TYPE: PRT
, ORGANISM: Homo sapiens
US-60-659-397-1215

```

[illegible]

```

RESULT 11
US-60-659-397-1215
; Sequence 1215, Application US/60659397
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: CHANG, Sheng-Yung
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RESPONSE TO INTERFERON TREATMENT IN HEPATITIS C
; TITLE OF INVENTION: VIRUS-INFECTED SUBJECTS, METHODS OF DETECTION AND
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001470
; CURRENT APPLICATION NUMBER: US/60/659,397
; CURRENT FILING DATE: 2005-03-09
; NUMBER OF SEQ ID NOS: 47859
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 1215
; LENGTH: 1195
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-659-397-1215

```


Db 268 SVSEDLLEFFIAVSAQAWNHQSGVVRPDSHSQSGP-----NSDPGLEPED-----SNSTSA 318
QY 357 DSQP-----QSKSKNTWYIDEVADPAKSLTE-----ISTDFRSSPPLQ- 396
Db 319 LEDPLEFLDMABTEKICDYLFNVSQSSALNAKNGLTAKARDINAVLIDMERQGVYRQ 378
QY 397 --PPVNSLTENR-----FHSLP-----FSLTKMPNTNGS-----IGH 428
Db 379 GTTPPIWHLTDKKRERMOIKRNTNSVPETAPAAIPETRRNAEFLTCNIPTSNASNNMVT 438
QY 429 SPLSLSAQSVMBELNAPVQESPLAMPNGSHGLEVGLAEVKENPPFVGVIRWIGQ-- 486
Db 439 EKVENQEPVIKLENQEARPEPARLKPVPVHYNGPSKAGVDFENG-----QWATDDI 491
QY 487 PPGLEVLAGLELEDEACAGCTDGTFRGTRYFTCALKKALFKV--LKSCRDPDSRFASLQ-- 542
Db 492 PDLNSIRA-----APGEFR-----AIMEPSPFYSHGLPRCSYKYLTECOLK 534
QY 543 -PVSNOIERCNSLAFGYLSE--VVENTPPKMEKEGLEIMIGKKGIQGHYNSCYLDST 599
Db 535 NPISGLLEYAQ---FASQTCFNMIEQSGPPHPRFKFQVINGREFPPPAEAGSKKVAQ 591
QY 600 LFCLEAFSSVLDVLLRPKEKN-DVEYVYSETQELLRT 635
Db 592 DAAMKANTILLEAKAKDSKSEESHYSTEKESKT 628

RESULT 14

US-60-659-397-1214
; Sequence 1214, Application US/60659397
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: CHANG, Sheng-Yung
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RESPONSE TO INTERPERON TREATMENT IN HEPATITIS C
; TITLE OF INVENTION: VIRUS-INFECTED SUBJECTS, METHODS OF DETECTION AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001470
; CURRENT APPLICATION NUMBER: US/60/659,397
; NUMBER OF SEQ ID NOS: 47859
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1214
; LENGTH: 1213
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-659-397-1214

Query Match 2.2%; Score 113; DB 8; Length 1213;
Best Local Similarity 19.9%; Pred. No. 11;
Matches 115; Conservative 81; Mismatches 249; Indels 132; Gaps 26;
QY 142 PGVVRERGP-LLAERT-----VSGIFFGVELLEEGRGQGTGCVYQKOLFQCDSDCGF 194
Db 101 PG-LRPRFPVLLASSTRGRQVDIRGVRGVHLGSGQLGRGFQHPSPRGRSLPQRGVDCLS 159
QY 195 VALDKLELIEDDDTA-----LESYAGPGDTM-----QVELPPLLEINS---RVSLKGETI 242
Db 160 SHFQELSIYQDQQRILKLEELGEGKATTAHDLGKLTGPKKEINRVLSLAKKGLQK 219
QY 243 ESGTVIFCDVLPKESLGIVGVDMNDPIGNWDRGDFGVLCSPACVESTILLHINDIPE 302
Db 220 EAGTPPLMKIAVSTQAWNHQSGVVRP-----DGHSGQAFNSDPSLEP-----EDRNST 267
QY 303 SVTQERRPPKLAEMSR-----GVGDKGSSHNKPKATGSTDPGNRRSELFTYLNSSV 356
Db 268 SVSEDLLEFFIAVSAQAWNHQSGVVRPDSHSQSGP-----NSDPGLEPED-----SNSTSA 318
QY 357 DSQP-----QSKSKNTWYIDEVADPAKSLTE-----ISTDFRSSPPLQ- 396
Db 319 LEDPLEFLDMABTEKICDYLFNVSQSSALNAKNGLTAKARDINAVLIDMERQGVYRQ 378
QY 397 --PPVNSLTENR-----FHSLP-----FSLTKMPNTNGS-----IGH 428

Db 379 GTTPPIWHLTDKKRERMOIKRNTNSVPETAPAAIPETRRNAEFLTCNIPTSNASNNMVT 438
QY 429 SPLSLSAQSVMBELNAPVQESPLAMPNGSHGLEVGLAEVKENPPFVGVIRWIGQ-- 486
Db 439 EKVENQEPVIKLENQEARPEPARLKPVPVHYNGPSKAGVDFENG-----QWATDDI 491
QY 487 PPGLEVLAGLELEDEACAGCTDGTFRGTRYFTCALKKALFKV--LKSCRDPDSRFASLQ-- 542
Db 492 PDLNSIRA-----APGEFR-----AIMEPSPFYSHGLPRCSYKYLTECOLK 534
QY 543 -PVSNOIERCNSLAFGYLSE--VVENTPPKMEKEGLEIMIGKKGIQGHYNSCYLDST 599
Db 535 NPISGLLEYAQ---FASQTCFNMIEQSGPPHPRFKFQVINGREFPPPAEAGSKKVAQ 591
QY 600 LFCLEAFSSVLDVLLRPKEKN-DVEYVYSETQELLRT 635
Db 592 DAAMKANTILLEAKAKDSKSEESHYSTEKESKT 628

RESULT 15

US-11-033-545-313
; Sequence 313, Application US/11033545
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/11/033,545
; CURRENT FILING DATE: 2005-01-12
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 313
; LENGTH: 1226
; TYPE: PRT
; ORGANISM: Human
US-11-033-545-313

Query Match 2.2%; Score 113; DB 7; Length 1226;
Best Local Similarity 19.9%; Pred. No. 11;
Matches 115; Conservative 81; Mismatches 249; Indels 132; Gaps 26;
QY 142 PGVVRERGP-LLAERT-----VSGIFFGVELLEEGRGQGTGCVYQKOLFQCDSDCGF 194
Db 69 PG-LRPRFPVLLASSTRGRQVDIRGVRGVHLGSGQLGRGFQHPSPRGRSLPQRGVDCLS 127
QY 195 VALDKLELIEDDDTA-----LESYAGPGDTM-----QVELPPLLEINS---RVSLKGETI 242
Db 128 SHFQELSIYQDQQRILKLEELGEGKATTAHDLGKLTGPKKEINRVLSLAKKGLQK 187
QY 243 ESGTVIFCDVLPKESLGIVGVDMNDPIGNWDRGDFGVLCSPACVESTILLHINDIPE 302
Db 188 EAGTPPLMKIAVSTQAWNHQSGVVRP-----DGHSGQAFNSDPSLEP-----EDRNST 235
QY 303 SVTQERRPPKLAEMSR-----GVGDKGSSHNKPKATGSTDPGNRRSELFTYLNSSV 356
Db 236 SVSEDLLEFFIAVSAQAWNHQSGVVRPDSHSQSGP-----NSDPGLEPED-----SNSTSA 286
QY 357 DSQP-----QSKSKNTWYIDEVADPAKSLTE-----ISTDFRSSPPLQ- 396
Db 287 LEDPLEFLDMABTEKICDYLFNVSQSSALNAKNGLTAKARDINAVLIDMERQGVYRQ 346
QY 397 --PPVNSLTENR-----FHSLP-----FSLTKMPNTNGS-----IGH 428
Db 347 GTTPPIWHLTDKKRERMOIKRNTNSVPETAPAAIPETRRNAEFLTCNIPTSNASNNMVT 406
QY 429 SPLSLSAQSVMBELNAPVQESPLAMPNGSHGLEVGLAEVKENPPFVGVIRWIGQ-- 486
Db 407 EKVENQEPVIKLENQEARPEPARLKPVPVHYNGPSKAGVDFENG-----QWATDDI 459

```
QY 487 PPGNEVLAGELEDEACGCTDGTFRGTRYFTCALKKALFVK--LKSCRPDSRPASLQ-- 542
Db 460 PDDLNSIRA-----APGEFR-----AIMEMPSYSHGLPRCSPYKXLTECQLK 502
QY 543 -PVSNOIERCNSLAFGGYLSE--VVEENTPPKMEKEGLEIMIGKKGIQGHYNSCYLDST 599
Db 503 NPISGLLEVAQ---FASQTCEFNMIQSGPPHPRFKFQVINGREPPPAEAGSKKVAQ 559
QY 600 LFCLFAPSSVLDTVLLRPKEKN-DVEYYSETQELLRT 635
Db 560 DAAMKAMTILLEAKADSGKSESSHYTEKESEKT 596
```

Search completed: April 18, 2005, 15:33:47
Job time : 58 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 18, 2005, 15:55:11 ; Search time 592 Seconds
(without alignments)
2496.870 Million cell updates/sec

Title: US-09-671-687A-3

Perfect score: 5034

Sequence: 1 MSSGLWSQKVTSPWEERI.....RLLCDAYCMYQSPMTSLYK 949

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9157160 seqs, 778792062 residues

Total number of hits satisfying chosen parameters: 18314320

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgn2_1/USPTO_spool_p/US09671687/runat 18042005 115020 17215/app query.fasta_1.1095
-DB=Pending Patents NA_New -QPMT=fastap -SUFFIX=rnnpn -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09671687 @CGN 1 1 268 @runat 18042005 115020 17215 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:
2: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq2:
3: /cgn2_6/ptodata/1/pna/US07_NEW_COMB.seq:
4: /cgn2_6/ptodata/1/pna/US07_NEW_COMB.seq2:
5: /cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq:
6: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:
7: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:
8: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq2:
9: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq:
10: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq2:
11: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq3:
12: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4971.5	98.8	5371	7	US-10-760-678-3 Sequence 3, Appli
2	3676	73.0	2523	7	US-10-122-851-290 Sequence 290, App
C 3	1490.5	29.6	1160	8	US-10-450-763-26238 Sequence 26238, A
C 4	239	4.7	600	7	US-10-972-079-20538 Sequence 20538, A
C 5	234.5	4.7	600	7	US-10-972-079-20539 Sequence 20539, A
C 6	234.5	4.7	600	7	US-10-972-079-20540 Sequence 20540, A

7	160.5	3.2	4847	10	US-11-036-196-1764	Sequence 1764, Ap
8	151	3.0	5793	12	US-60-660-557-55	Sequence 55, Appl
9	151	3.0	5898	12	US-60-660-557-54	Sequence 54, Appl
10	146	2.9	5449	2	PCT-US05-07748-423	Sequence 423, App
11	146	2.9	5554	2	PCT-US05-07748-422	Sequence 422, App
12	146	2.9	5555	2	PCT-US05-07748-705	Sequence 705, App
13	136	2.7	2154	7	US-10-932-182A-77607	Sequence 77607, A
14	128	2.5	1497	8	US-10-450-763-21942	Sequence 21942, A
C 15	127	2.5	3918	7	US-10-467-657A-5405	Sequence 5405, Ap
C 16	127	2.5	3918	7	US-10-467-657A-5405	Sequence 5405, Ap
17	126	2.5	4314	12	US-60-651-509-1122	Sequence 1122, Ap
18	126	2.5	4314	12	US-60-651-509-1122	Sequence 1122, Ap
19	126	2.5	4314	12	US-60-651-509-1122	Sequence 1122, Ap
20	126	2.5	4314	12	US-60-651-509-1122	Sequence 1122, Ap
21	126	2.5	6521	12	US-60-651-509-1123	Sequence 1123, Ap
22	126	2.5	6521	12	US-60-651-509-1123	Sequence 1123, Ap
C 23	125.5	2.5	4374	8	US-10-450-763-21946	Sequence 21946, A
24	123.5	2.5	4249	7	US-10-489-448-399	Sequence 399, App
25	123.5	2.5	5810	12	US-60-660-590-101	Sequence 101, App
26	123.5	2.5	5894	1	PCT-US05-10257-421	Sequence 421, App
27	123.5	2.5	7789	7	US-10-696-909A-45	Sequence 45, Appl
28	123.5	2.5	9027	7	US-10-696-909A-43	Sequence 43, Appl
29	123	2.4	1506	7	US-10-932-182A-2930	Sequence 2930, Ap
C 30	122.5	2.4	31826	10	US-11-031-175-1256	Sequence 1256, Ap
31	122	2.4	1500	7	US-10-932-182A-76457	Sequence 76457, A
C 32	120	2.4	2089378	7	US-10-526-324-1	Sequence 1, Appli
C 33	120	2.4	2089378	7	US-10-526-324-342	Sequence 342, App
C 34	120	2.4	2089378	7	US-10-526-324-723	Sequence 723, App
35	120	2.4	2089378	7	US-10-526-324-1087	Sequence 1087, Ap
36	120	2.4	2089378	7	US-10-526-324-1459	Sequence 1459, Ap
37	120	2.4	2089378	7	US-10-526-324-1838	Sequence 1838, Ap
C 38	119.5	2.4	5715	1	PCT-US05-10257-236	Sequence 236, App
C 39	119.5	2.4	9359	8	US-10-450-763-22954	Sequence 22954, A
40	119	2.4	6343	12	US-60-659-397-310	Sequence 310, App
41	119	2.4	6487	12	US-60-659-397-309	Sequence 309, App
42	119	2.4	11058	10	US-11-032-569-1	Sequence 1, Appli
C 43	118	2.3	2253	7	US-10-932-182A-79868	Sequence 79868, A
C 44	118	2.3	16047	10	US-11-031-175-1136	Sequence 1136, Ap
45	117.5	2.3	1828	2	PCT-US05-02325-51	Sequence 51, Appl

ALIGNMENTS

RESULT 1

US-10-760-678-3
; Sequence 3, Application US/10760678

; GENERAL INFORMATION:

; APPLICANT: Derry, Jonathan

; APPLICANT: Fanslow, William

; APPLICANT: Dougall, William

; TITLE OF INVENTION: SCREENING ASSAYS FOR AGONISTS OR ANTAGONISTS OF CD40 SIGNALING

; FILE REFERENCE: 3198

; CURRENT APPLICATION NUMBER: US/10760,678

; CURRENT FILING DATE: 2004-01-20

; PRIOR APPLICATION NUMBER: US/09/851,673

; PRIOR FILING DATE: 2001-05-08

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: Patentin version 3.0

; SEQ ID NO 3

; LENGTH: 5371

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (392)..(3262)

; US-10-760-678-3

Alignment Scores:

Pred. No.: 0 Length: 5371
Score: 4971.50 Matches: 948
Percent Similarity: 99.16% Conservative: 0
Best Local Similarity: 99.16% Mismatches: 1
Query Match: 98.76% Indels: 7

DB:	7	Gaps:	5
US-09-671-687A-3 (1-949) x US-10-760-678-3 (1-5371)			
QY	1	MetSerSerGlyLeuTyrSerGlnGluLysValThrSerProTyrTrpGluGluArgIle	20
DB	392	ATGAGTTGAGGCTTATGGAGCCAGAAAAGATCACTTCCCTCTACTGGGAAGCGGATT	451
QY	21	PheTyrLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLeuLys	40
DB	452	TTTTACTTCTTCTTCAAGAATGCAGGTTACAGACAAACAAACAAAGCTCCTTAA	511
QY	41	ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisArgIle	60
DB	512	GTACCGAAGGGAAGTATAGACAGTATATTCAGATCGTTCGTGGGGCAATCAAGGATT	571
QY	61	ProSerAlaLysGlyLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla	80
DB	572	CCTTCTGCAAAAGGCAAGAAAATCAGATTGGATTAAATAATCTAGGCAACCTCATGCA	631
QY	81	ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu	99
DB	632	GTTCCTTTGTGATGAAGAAGGATGTTGTAGAGATAAATGAAAAGTTTCACAGATTACTT	691
QY	100	LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer	119
DB	692	TTGGCAATTACCAATTGTGAGGAGAGGTTTCAGCTGTTTAAAAACAGAAACAGACTAAGT	751
QY	120	LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu	139
DB	752	AAAGCGCTCCAAATAGACTGGGCTCTCTGTGAAAGTACAGCTGAGATCTGGGGAAGAA	811
QY	140	LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly	159
DB	812	AAATTTCTGGAGTTGTAGCTTCAGAGACCCCTGTTAGCAGAGAGGACAGTCTCCGGA	871
QY	160	IlePhePheGlyValGluLeuLeuGluGlyArgGlyGlnGlyPheThrAspGlyVal	179
DB	872	ATATTCTTTGGAGTTGAAATTGCTGGAAGAAGGTCGTGGTCAAGTTTCACTGACGGGGTG	931
QY	180	TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp	198
DB	932	TACCAAGGAAACAGCTTTTTCAGTGTGATGAAGATTGTGGCGTGTGTTGTCATTGGAC	991
QY	199	LysLeuLeuLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly	218
DB	992	AAGCTAGAACTCATAGAAGATGATGACTGCATTGGAAGTGAATTACGAGGCTCTCGG	1051
QY	219	AspThrMetGlnValGluLeuProLeuGluIleAsnSerArgValSerLeuLysGly	238
DB	1052	GACACAATGCAGGTCGAACTTCTCTCTTTGGAAATAAACTCCAGAGTTCCTTTGAAGTT	1111
QY	239	GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer	258
DB	1112	CGAGAAACAATAAGAACTCGAAACAGTTATATCTGTGATGTTTGGCAGGAAAGAAAGC	1171
QY	259	LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe	278
DB	1172	TTAGGATATTTTGTGGTGGACATGGATAACCCCTATTGGCAACTGGGATGGAAGATT	1231
QY	279	AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn	297
DB	1232	GATGGAGTCAGCTTTGTAGTTTGTGTGTGTTGAAAGTACAAATTCATTGTCACATCAAT	1291
QY	298	AspIleIlePro-----GluSerValThrGlnGluArgProProLysLeuAla	314
DB	1292	GATATCATCCAGCTTTATCAGAGAGGTGTGACGAGAAAGGAGGCTCCCAAACCTGGC	1351
QY	315	PheMetSerArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThr	334
DB	1352	TTTATGTCAGAGGTTGTGGGCAAGGTTTCATCCAGTCATATAAACAAGAGGCTACA	1411
QY	335	GlySerThrSerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGly	353
DB	1412	GGATCTACCTCAGACCTCGAAATAGAAACAGATCTGAAATATTTTATATACCTTAAATGGG	1471
QY	354	SerSerValAspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGluVal	373
DB	1472	TCTTCTGTGTGACTCACACACACATCCAAATCMAAAAATACATGGTACATTTGATGAAGTT	1531
QY	374	AlaGluAspProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerPro	393
DB	1532	GCAGAAGACCTCGAAAATCTCTTACAGAGATATCTACAGACTTTGACCGTTCTTACCA	1591
QY	394	ProLeuGlnProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuPro	413
DB	1592	CCACTCCAGCCTCCTCTGTGAACCTCACTGACACCCAGAACAGATTCCACTCTTTACCA	1651
QY	414	PheSerLeuThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeu	433
DB	1652	TTCAGTCTCACCAAGATGCCAATACCAATGGAAGTATTGGCCACAGTCCACTTCTCTG	1711
QY	434	SerAlaGlnSerValMetGluGluLeuAsnThrAlaProValGlnGluSerProLeu	453
DB	1712	TCAGCCCACTGTATATGGAAGAGCTAAACACTGCACCCGTCACAGAGAGTCCACCCCTG	1771
QY	454	AlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGlu	473
DB	1772	GCCATGCTCTCGGGAACCTCACATGCTAGAAAGTGGGCTCATTTGGCTGAAGTTAAGGAG	1831
QY	474	AsnProProPheTyrGlyValIleArgTyrIleGlyGlnProProGlyLeuAsnGluVal	493
DB	1832	AACCTCTCTTCTATGGGGTAAATCCGTGGATCGGCAGCCACAGAGACTGAATGAAGTG	1891
QY	494	LeuAlaGlyLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGly	513
DB	1892	CTCGCTGGACTGGAACCTGGAAGATGAGTGTGACGGCTGTACGGATGGAACTTCAGAGGC	1951
QY	514	ThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArg	533
DB	1952	ACTCGGTAATTCACCTGTGCTCGCTGGAAGAAGCGCTGTTGTGAAACTGAAGAGCTG	2011
QY	534	ProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSer	553
DB	2012	CCTGACTTAGGTTTCATCTGACCGCGGTTTCCAATCAGATTGAGCGCTGTAATCTCT	2071
QY	554	LeuAlaPheGlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGlu	573
DB	2072	TTAGCATTTGGAGCTACTTTAAGTGAAGTAGTAGAAGAAATACTCCACCAAAATGGAA	2131
QY	574	LysGluGlyLeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSer	593
DB	2132	AAAGAAGGCTTGGAGATAATGATTGGGAAGAAAGGCATCCAGGGTCAATCAATCTCT	2191
QY	594	CysTyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrVal	613
DB	2192	TGTTACTTAGACTCAACCTTATTTCTGCTTATTTTGTCTTTTAGTCTGTGTCGACACTGTG	2251
QY	614	LeuLeuArgProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeu	633
DB	2252	TTACTTAGACCCCAAGAAAGAACGATGTAGAATATTATATAGTGAACCCAGAGACTACTG	2311
QY	634	ArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMet	653
DB	2312	AGGACAGAAATTTGTTAATCTCTGAGATAATATGATATGTGTGTGTCACCAAAATATG	2371
QY	654	LysLeuArgLysIleLeuGluLysValGluAlaAspSerGlyPheThrSerGluGluLys	673
DB	2372	AAACTGAGGAAAACTTGAAGAGGTGGAGCTGCATCAGGATTTTACCTCTCGAAGAAAA	2431
QY	674	AspProGluGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeu	693
DB	2432	GATCCTGAGGAATCTTGAATATTTCTGTTTTCATCATATTTTAAGGGTAGAACCTTTGCTA	2491
QY	694	LysIleArgSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGlu	713

```
Db 2492 AAAATAAGATCAGCAGGTCAAAAGGTACAAGATTGTTTACTTCTATCAAAATTTTATGGAA 2551
QY 714 LysAsnGluLysValGlyValProThrIleGlnLeuLeuGluTrpSerPheIleAsn 733
Db 2552 AAAAATCAGAAATGGCGTTCACCAATTCAGCAGTGTGTAGAAATGGTCTTTTATCAAC 2611
QY 734 SerAsnLeuLysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGly 753
Db 2612 AGTAACCTGAAATTTGAGAGGACCACATCATGCTGATTATTCAGATGCGCTCGATTGGGA 2671
QY 754 LysAspPheLysLeuPheLysIlePheProSerLeuLeuLeuAsnIleThrAspLeu 773
Db 2672 AAAGACTTTAAACTATTAAAAAATTTTCTCTCTGGAATTAATAACAGATTGA 2731
QY 774 LeuGluAspThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArg 793
Db 2732 CTTGAAGACATCCAGACAGTCCGATATGTGGAGGGCTTGCATATGTATGAGTGTAGA 2791
QY 794 GluCysTyrAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCys 813
Db 2792 GAATGCTACAGCATCCGACATCTCAGCTGGAAAAATCAAGCAGTTCGTAAACCTGC 2851
QY 814 AsnThrGlnValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeu 833
Db 2852 AACACTCAAGTCCACCTTCATCCGAGAGGCTGATCATTAATATAACCCAGTGTCACTT 2911
QY 834 ProLysAspLeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGlu 853
Db 2912 CCCAAAGACTTACCCGACTGGGACTGGAGACACGGCTGCATCCCTTGGCCAGATATGGAG 2971
QY 854 LeuPheAlaValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLys 873
Db 2972 TTATTTGCTGTTCTCTGCATAGAAACAAAGCCACTATGTTGCTTTTGTGAAGTATGGGAG 3031
QY 874 AspAspSerAlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGly 893
Db 3032 GACGATTCGCCCTGGCTCTCTTTGACAGCATGCCCATCGGATGTTGTGAGATGGC 3091
QY 894 PheAsnIleProGlnValThrProCysProGluValGlyGlyTyrLeuLysMetSerLeu 913
Db 3092 TTCAACATTCCTCAAGTCAACCCATGCCAGAGAGTAGGAGACTTTGAAGATGTCCTG 3151
QY 914 GluAspLeuHisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCys 933
Db 3152 GAAGACCTGCATTCCTTGGACTCCAGGAGAAATCCAGGCTGTGCACGAGACTGCTTTGT 3211
QY 934 AspAlaTyrMetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
Db 3212 GATGCTATATGTGTCATGTACAGAGTCCAAACAATGAGTTTGTACAAA 3259
```

RESULT 2

```
US-10-122-851-290
; Sequence 290, Application US/10122851
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyang
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and
; FILE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2BDV3
; CURRENT FILING DATE: US/10/122,851
; PRIOR APPLICATION NUMBER: 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt FL_genes Version 1.0
; SEQ ID NO 290
```

```
; LENGTH: 2523
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (81)..(2138)
US-10-122-851-290
Alignment Scores:
Pred. No.: 3,03e-242 Length: 2523
Score: 3676.00 Matches: 692
Percent Similarity: 99.28% Conservative: 0
Best Local Similarity: 99.28% Mismatches: 3
Query Match: 73.02% Indels: 2
DB: 7 Gaps: 2
US-09-671-687A-3 (1-949) x US-10-122-851-290 (1-2523)
```

```
QY 255 GlyLysGluSerLeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrp 274
Db 45 GGAATAATAAGAGCTTTTATATATTTTGGTGGTGGACATGGATAACCCCTATTGGCAACTGG 104
QY 275 AspGlyArgPheAspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeu 293
Db 105 GATGGAAGATTGATGGAGTGCAGCTTTGTAGTTTTCGTGTGTGTAAGATACAATTCTA 164
QY 294 LeuHisIleAsnAspIleIleProGluSerValThrGlnGluArgArgProProLysLeu 313
Db 165 TTGCACATCAATGATATATATCCAGAGAGTGTGACGAGAAAGAGAGGCTCCCAAACTT 224
QY 314 AlaPheMetSerArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAla 333
Db 225 GCCTTTATGTCAAGAGGTGTTGGGACAAAAGTTTCATCCAGTCAATAATAAACCAAGGCT 284
QY 334 ThrGlySerThrSerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsn 352
Db 285 ACAGGATCTACCTCAGACCCCTGGAAATAGAAACAGATCTGAATTTATTTATACCTTAAAT 344
QY 353 GlySerSerValAspSerGlnProGlnSerLysSerLysAsnThrTyrTyrIleAspGlu 372
Db 345 GGGTCTCTGTTGACTCAGAACACACATCCAAATCAAAAATAACATGATGATGATGAA 404
QY 373 ValAlaGluAspProAlaLysSerLeuThrGluLeSerThrAspPheAspArgSerSer 392
Db 405 GTTGCAAGAGACCCCTGCAAAATCTCTTACAGAGATATCTACAGACTTTTGACCGTTCTTCA 464
QY 393 ProProLeuGlnProProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeu 412
Db 465 CCACCACTCCAGCCTCTCTCTGTGAACTCACTGACCCAGAGAACAGATTTCCACTCTTTA 524
QY 413 ProPheSerLeuThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSer 432
Db 525 CCATTTCAGTCTCACAAGATGCCCAATACCAATGGAAGTATTGGCCACAGTCCACTTCT 584
QY 433 LeuSerAlaGlnSerValMetGluLeuLeuAsnThrAlaProValGlnGluSerProPro 452
Db 585 CTGTCAGCCAGCTCTGTAATGGAAGAGCTAAACACTGCACCCGTCACAGAGAGTCCACCC 644
QY 453 LeuAlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLys 472
Db 645 TTGGCCATGGCTCTCTGGGAACTCAGATGGTCTAGAAAGTGGGCTCATTTGGCTGAAGTTAAG 704
QY 473 GluAsnProPheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGlu 492
Db 705 GAGAACCCCTCTTCTATGGGGTAAATCCGTGGATCGGTGACCCACAGGACTTGAATGAA 764
QY 493 ValLeuAlaGlyLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArg 512
Db 765 GTGCTCGCTGGACTGGAACTGGAAAGATGAGTGTCCAGGCTGTACGGATGGAACTTCA 824
QY 513 GlyThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCys 532
Db 825 GGCACCTGGTATTTCACCTGTGCCCTCGAAGAGCGCTGTTTGTGAAACTGAAGAGCTGC 884
```

```
QY 533 ArgProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsn 552
DB 885 AGCCCTGACTCTAGGTTTGCATCATTTGCAGCCGGTTTCCAAATCAGATTGAGCGCTGTAAAC 944
QY 553 SerLeuAlaPheGlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMet 572
DB 945 TCTTTAGCATTTGGAGGCTACTTAAGTGAAGTAGTAGAAGAAATACTCCACAAAATG 1004
QY 573 GluLysGluGlyLeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsn 592
DB 1005 GAAAAGAAGGCTTGGAGATATGATTGGAGAGAGAAAGGCATCCAGGGTCATTACAT 1064
QY 593 SerCysTyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThr 612
DB 1065 TCTTGTTACTTAGACTCAACCTTATTCGCTTATTTGCTTTTAGTTCTGTTCTGGACACT 1124
QY 613 ValLeuLeuArgProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeu 632
DB 1125 GTGTTACTTAGACCCCAAGAAAGAACGATGTAGAAATATTATAGTGAAACCCCAAGAGCTA 1184
QY 633 LeuArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIle 652
DB 1185 CTGAGGACGAAATTTGTTAATCTCTGAGATATATGGATATGTGTGCCACAAAATTT 1244
QY 653 MetLysLeuArgLysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGlu 672
DB 1245 ATGAAACTGAGGAAATACTTGAAGAGGTGGAGGCTGCATCAGGATTTACCTCTGAAGAA 1304
QY 673 LysAspProGluGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeu 692
DB 1305 AAAGATCTCTGAGGAATCTTGAATATCTGTTTCATCATATTTAAGGGTAGAACCTTTG 1364
QY 693 LeuLysIleArgSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMet 712
DB 1365 CTAAATAAAGATCAGCAGGTCAAAAGGTACAGATTGTTACTTCTATCAAAATTTTATG 1424
QY 713 GluLysAsnGluLysValGlyValProThrIleGlnGlnLeuLeuGluTyrSerPheIle 732
DB 1425 GAAAAAATGAGAAAGTTGGCGTTCACCAATTCAGCAGTTGTAGATGGTCTTTATC 1484
QY 733 AsnSerAsnLeuLysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPhe 752
DB 1485 AACAGTAACCTGAAATTTGCAGAGGCACCATCATGTCTGATTTATTCAGATGCCCTCATTT 1544
QY 753 GlyLysAspPheLysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAsp 772
DB 1545 GGAAGAGACTTTAAACATATTTAAAAAAATTTTCTCTCTCGAATTTAAATATAACAGAT 1604
QY 773 LeuLeuGluAspThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCys 792
DB 1605 TTACTTTGAAGACACTCCAGACAGTCCCGATATGTGGAGGGCTTGCATATGATAGTGT 1664
QY 793 ArgGluCysTyrAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThr 812
DB 1665 AGAGATGCTACGACGATCCGACATCTCAGCTGGAAAAATCAAGCAGTTTGTGAAAAACC 1724
QY 813 CysAsnThrGlnValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSer 832
DB 1725 TGCAACACTCAAGTCCACCTTCTATCCGAAGAGGCTGAATCATAAATATATACCCAGGTGCA 1784
QY 833 LeuProLysAspLeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMet 852
DB 1785 CTTCCCAAGAGACTTACCCGACTGGGACTGGAGACAGCGTGCATCTCTTCCAGATATG 1844
QY 853 GluLeuPheAlaValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGly 872
DB 1845 GAGTTATTTGCTGCTCTCTGCATAGAAACAGCCACTATGTTGCTTTTGTGAAGTATGG 1904
QY 873 LysAspAspSerAlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsn 892
DB 1905 AAGGACGATTTCCCTGGCTCTCTTTTACAGCAGATGCCGATCGGATGTGTGTGAGAT 1964
```

```
QY 893 GlyPheAsnIleProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSer 912
DB 1965 GGCTTCAACATTCCTCAAGTCAACCCATGCCAGAAAGTAGGAGACTCTTGAAGATGCT 2024
QY 913 LeuGluAspLeuHisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeu 932
DB 2025 CTGGAAGACTGCAATTCCTTGGACTCCAGGAGAAATCCAGGCTGTGACCGAGACTGCTT 2084
QY 933 CysAspAlaTyrMetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
DB 2085 TGTGATGCATATATGTGCATGTACACAGAGTCCCAACAATGAGTTGTGACAA 2135

RESULT 3
US-10-450-763-26328/c
; Sequence 26328, Application US/10450763
; GENERAL INFORMATION:
; APPLICANT: Hveeq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 26328
; LENGTH: 1160
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (639)..(67)
; OTHER INFORMATION: 96% homologous to Homo sapiens Familial Cylindromatosis
; OTHER INFORMATION: Gene, accession number AJ250014, Smith-Waterman Score=963.
US-10-450-763-26328

Alignment Scores:
Pred. No.: 1,09e-92 Length: 1160
Score: 1490.50 Matches: 295
Percent Similarity: 97.38% Conservative: 2
Best Local Similarity: 96.72% Mismatches: 4
Query Match: 29.61% Indels: 4
DB: 8 Gaps: 3

US-09-671-687A-3 (1-949) x US-10-450-763-26328 (1-1160)

QY 1 MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluGluArgIle 20
DB 979 ATGAGTTTCAGGCTTATGGAGCCCAAGAAAGTCACTTCACCTACTCTGGGAGAGCGGATT 920
QY 21 PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLys 40
DB 919 TTTTACTTGTCTTCTCAAGAAATGCAGGTTACAGACAAACAAACAAAGCTCTCTTAA 860
QY 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHiserArgIle 60
DB 859 GTACCCGAAGGGAAGTAGTAGACAGTATATTCAGATCGTTCTGTGGGCGCATTCAGGATT 800
QY 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla 80
DB 799 CCTTCTGCAAAAGGCAAGAAATAATCAGATTGGATTTAAAAATTTAGAGCAACCTCATGCA 740
QY 81 ValLeuPheValAspGlu---AspValValGluIleAsnGlnLysPheThrGluLeuLeu 99
DB 739 GTTCTCTTGTGATGAAAGGATGTTGTAGAGATAAATGATATAAGTTCACAGAGTTACTT 680
QY 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArg-LeuSe 119
DB 679 TTGGCAATTACCAATTGTGAGGAGAGGTTTCAGCCTGTTTAAACCCAGAAACCCAGTTTAAG 620
```


QY 119 rlysglyleuGlnlleAspValglyCysProVallyysValGlnLeuArgSerGlyGluGI 139
Db 619 TAAAGCCCTCCAAATAGACGGGGCTGTCTCTGTGAAAGTACACTGAGATCTCGGGAAGA 560
QY 139 ulyysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGI 159
Db 559 AAAATTTCTCGAGTGTGACCTTCAGAGGACCCCTGTAGCAGAGAGACAGTCTCCGG 500
QY 159 yllePhePheGlyValGluLeuLeuGluGluArgGlyGlnGlnGlnPheThrAspGlyVa 179
Db 499 AATATTCTTTGGAGTGAATTTGCTGGAAGAAGGCTGCTGGTCAAGGTTTCACTGACGGGT 440
QY 179 lTyrGlnGlyysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAs 198
Db 439 GTACCAAGGGAACACACTTTTTCAGTGTGTGATGAAGATTGCGCGTGTGTTTGCATTGGA 380
QY 198 plysLeuGluLeuLeuGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGI 218
Db 379 CAAGCTAGAACTCATAGAAGATGATGACACTGCTATTGGAATGATTTACGAGTCTCGG 320
QY 218 yAspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGI 238
Db 319 GGACACAATGCAGTCTGAACTTCTCTCTTTTGGAAATAAACTCCAGAGTTCTTTGAAGT 260
QY 238 yGlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSe 258
Db 259 TGGAGAAACAATAGAACTTGGAACTGATATATCTGTGATGTTTGGCAGGAAAGAAAG 200
QY 258 rleuGlyTyrPheValGlyValAspMetAspAsnProIleGlyValAsnTrpAspGlyArgPh 278
Db 199 CTTAGCATATTGTTGGTGTGGACATGGATAACCCCTATTGGCAACTGGGATGGAGATT 140
QY 278 eAspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAs 297
Db 139 TGATGGAGTGCAGCTTTTGTAGTTTTCGCTGTGTGTAAGTACAAATCTATTGCACATCAA 80
QY 297 nAspIleIlePro 301
Db 79 TGATATCATCCCCA 67

RESULT 4
US-10-972-079-20538/c
; Sequence 20538, Application US/10972079
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: ROSENFELD, David
; APPLICANT: KERR, Richard
; APPLICANT: BATES, Stephen
; APPLICANT: HOLM, Tom
; TITLE OF INVENTION: METHODS & SYSTEMS FOR INFERRING TRAITS TO BREED & MANAGE NON-BEEF
; FILE REFERENCE: MM1110-2
; CURRENT APPLICATION NUMBER: US/10/972,079
; PRIOR FILING DATE: 2004-10-22
; PRIOR APPLICATION NUMBER: US 60/514,333
; NUMBER OF SEQ ID NOS: 96631
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20538
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Chicken 19866894231145_1
US-10-972-079-20538

Alignment Scores:
Pred. No.: 3,69e-07 Length: 600
Score: 239.00 Matches: 66
Percent Similarity: 50.64% Conservatives: 13
Best Local Similarity: 42.31% Mismatches: 47
Query Match: 4.75% Indels: 30
DB: 7 Gaps: 7

US-09-671-687A-3 (1-949) x US-10-972-079-20538 (1-600)
QY 696 ArgSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsn 715
Db 450 AGATCAGCAGGTCAAGAAATCAAGACTGTTATTTCTACCAATTTTATGACAAAAAT 391
QY 716 GluLysValGlyValProThrIleGlnGlnLeuLeuGluTrpSerPheIleAsnSerAsn 735
Db 390 GAGAAAGTTGGAGTCCCGACAATYCACAGTTACTGGAGTGGTCTCTTCATCAACAGCAAC 331
QY 736 LeuLysPheAlaGlu---AlaProSerCysLeuIleIleGlnMetProArgPheGlyLys 754
Db 330 TTGAAGTTTTCAGAGGTTGGTGACACTGKGTGTGTGTC-----AGAAATACTGAT 280
QY 755 AspPheLysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeu 774
Db 279 AGCTACATCTGTTTCAGC-----AATGAGAGAAAGTGGGA 244
QY 775 GluAspThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGlu 794
Db 243 TTGGGTGTGAAGGAAATGCAGA---TGTTTGGGATGGTGGAGTGGTCTCAGTG----- 193
QY 795 CysTyrAspAspProaspIleSerAlaGlyLysIleLysGlnPheCys----- 810
Db 192 -----LysTyrCysAsnThrGlnValHisLeuHisProLysArgLeu-----Asn 151
QY 811 -----LysTyrCysAsnThrGlnValHisLeuHisProLysArgLeu-----Asn 825
Db 150 GGTGTGGGTGTACCAAAATGCTTGTGATCTATTAGCCAGGAATTAATCAATAAANAAC 91
QY 826 HisLysTyrAsnProValSerLeuProLysAspLeuProAspTrpAsp 841
Db 90 TCTTCGTTACTGTTGTAAGTTGCCITTAAGATCTAGTATTTCGGAT 43

RESULT 5
US-10-972-079-20539/c
; Sequence 20539, Application US/10972079
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: ROSENFELD, David
; APPLICANT: KERR, Richard
; APPLICANT: BATES, Stephen
; APPLICANT: HOLM, Tom
; TITLE OF INVENTION: METHODS & SYSTEMS FOR INFERRING TRAITS TO BREED & MANAGE NON-BEEF
; FILE REFERENCE: MM1110-2
; CURRENT APPLICATION NUMBER: US/10/972,079
; PRIOR FILING DATE: 2004-10-22
; PRIOR APPLICATION NUMBER: US 60/514,333
; NUMBER OF SEQ ID NOS: 96631
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20539
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Chicken 19866894231145_2
US-10-972-079-20539

Alignment Scores:
Pred. No.: 7,51e-07 Length: 600
Score: 234.50 Matches: 45
Percent Similarity: 90.74% Conservatives: 4
Best Local Similarity: 83.33% Mismatches: 4
Query Match: 4.66% Indels: 1
DB: 7 Gaps: 1

US-09-671-687A-3 (1-949) x US-10-972-079-20539 (1-600)
QY 696 ArgSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsn 715
Db 383 AGATCAGCAGGTCAAGAAATCAAGACTGTTATTTCTACCAATTTTATGACAAAAAT 324

716	Qy	Glu	Y	Val	G	Val	Pro	Thr	Ile	G	nd	nd	nt	Leu	Leu	G	lu	Trp	Ser	Phe	Ile	Asn	Ser	Asn	735
323	Db	G	A	A	A	G	T	T	G	A	G	T	T	C	A	A	T	T	T	T	C	A	A	C	264
736	Qy	Leu	Y	S	Phe	Ala	G	lu	---	Ala	Pro	Ser	Cys	Leu	Ile	Ile	G	ln	748						
263	Db	T	T	G	A	G	T	T	T	G	C	A	A	C	T	T	G	C	T	G	T	C	A	222	

RESULT 6

```

US-10-972-079-20540/c
; Sequence 20540, Application US/10972079
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: ROSENFELD, David
; APPLICANT: KERR, Richard
; APPLICANT: BATES, Stephen
; APPLICANT: HOLM, Tom
; TITLE OF INVENTION: METHODS & SYSTEMS FOR INFERRING TRAITS TO BREED & MANAGE NON-BEEF
; TITLE OF INVENTION: LIVESTOCK
; FILE REFERENCE: MM1110-2
; CURRENT APPLICATION NUMBER: US/10/972,079
; CURRENT FILING DATE: 2004-10-22
; PRIOR APPLICATION NUMBER: US 60/514,333
; PRIOR FILING DATE: 2003-10-24
; NUMBER OF SEQ ID NOS: 96631
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 20540
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Chicken 19866894231145_3
US-10-972-079-20540

```

Alignment Scores:	
Pred. No.:	7,51e-07
Score:	234.50
Percent Similarity:	90.74%
Best Local Similarity:	83.33%
Query Match:	4.66%
DB:	7
Length:	600
Matches:	45
Conservative:	4
Mismatches:	4
Indels:	1
Gaps:	1

US-09-671-687A-3 (1-949) x US-10-972-079-20540 (1-600)

Qy	696	ArgSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsn	715
Db	320	AGATCAGCAGGTGCAGAAAGTRCAAGACTGTTATTCTTACCAAAATTTTATGCACAAAAT	261
Qy	716	GluLysValGlyValProThrIleGlnLeuLeuGluTyrSerPheIleAsnSerAsn	735
Db	260	GAGAAAGTTGGAGTCCCGACAAATYCAGCAGTTACTGGAGTGGTCTTCATCAACAGCAAC	201
Qy	736	LeuLysPheAlaGlu---AlaProSerCysLeuIleIleGln	748
Db	200	TTGAAGTTTGCAGAGGTGGTCACAACTGCTKGCTGTGTGTCAGA	159

RESINT 7

US-11-036-196-1764
 ; Sequence 1764, Application US/11036196
 ; GENERAL INFORMATION:
 ; APPLICANT: Mendrick, Donna
 ; APPLICANT: Porter, Mark
 ; APPLICANT: Johnson, Kory
 ; APPLICANT: Higgs, Brandon
 ; APPLICANT: Castle, Arthur
 ; APPLICANT: Elashoff, Michael
 ; TITLE OF INVENTION: Molecular Toxicology Modeling
 ; FILE REFERENCE: 44921-5089-US
 ; CURRENT APPLICATION NUMBER: US/11/036,196
 ; CURRENT FILING DATE: 2005-01-18
 ; PRIOR APPLICATION NUMBER: US/10/152,319
 ; PRIOR FILING DATE: 2002-05-22
 ; PRIOR APPLICATION NUMBER: US 60/292,335

```

; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/297,523
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,925
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,810
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/303,807
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/303,808
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/315,047
; PRIOR FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: US 60/324,928
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/330,867
; PRIOR FILING DATE: 2001-11-01
; Remaining prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2221

```

Alignment Scores:			
Pred. No.:	1.2	Length:	4847
Score:	160.50	Matches:	93
Percent Similarity:	32.76%	Conservative:	40
Best Local Similarity:	22.91%	Mismatches:	143
Query Match:	3.19%	Indels:	130
DB:	10	Gaps:	16

US-09-671-687A-3 (1-949) x US-11-036-196-1764 (1-4847)

301	Qy	ProGluSerValThrGlnGluArg-----ArgProProLysLeuAlaPheMet	316
272	Db	CCTAGTGGCATCACCATGCAGACCCAGTGGCTGAAGCCCCT-----	316
317	Qy	SerArgGlyValGlyAspLysGlySerSerSerHisAsnLysProLysAlaThrGlySer	336
317	Db	-----GGCCGAGGGGGAACACTCCAGTCCCGTGGTGAAGCCCTCCATC-----	361
337	Qy	ThrSerAspProGlyAsnArgArgSerGluLeuPheTyrThrLeuAsnGlySerSerVal	356
362	Db	-----GGGTCTGCCTCATCTTCTGTGTGGCATCGCCGAGCGGCTCCAAAGAA	409
357	Qy	AspSerGln-----ProGlnSerLysSerLysAsnThrTrpTyr	369
410	Db	GGGTCTCCCTGCACAACAGCGCTCGGGCCCTCTCGGTGGAGCCACCACT-----	463
370	Qy	IleAspGluValAlaGluAspProAlaLysSerLeuThrGluIleSerThrAspPhe---	388
464	Db	-----ACTGTCTCCGAGAACCAGGTCCAAAGGCAGCCGAAGTGGGTGATGACTTCTTG	517
389	Qy	-----AspArgSerSerProProLeuGlnProProValAsnSer	402
518	Db	GGAGACTTCGTGGTGGCGAGAGGGTGTGGGTGAATGCGCTGAAGCCCGCGTGGTCAG	577
403	Qy	LeuThrThrGluAsnArgPhe-----	409
578	Db	TACCTGGGTGAGACACAGTTTGCCTCAGCCAGCCAGTGGGCGGGTGGTCTTGATGATCTCT	637
409	Qy	-----	409
638	Db	GTGGGCAAGACGATGGAGCTGTGGTGGGTGCGCTACTTCGAGTGGCCACTCTACAG	697
410	Qy	-----HisSerLeuProPheSerLeuThrLysMetPro-----AsnThrAsnGlySer	425
698	Db	GGCATCTTCACAGCCCTCCAAAGTGAACCGCAGCCGAGCTCGAGGGCTCCCGCAGT	757

```
QY 426 IleGlyHisSerProLeuSerLeuSerAlaGlnSerValMetGluGluLeuAenThrAla 445
Db 758 GACGGCCACTCAGTGAATCCCTCACCGCCAGAACCTGCTCCTGCATTCGGCAGCGCC 817
QY 446 ProValGlnGluSerProLeu----- 453
Db 818 -----ACGCCCCCACTCACCGCGCGAGTCATCCCTCGGGAGAGTGTCTC 865
QY 454 -----AlaMetProProGlyAenSerHisGlyLeuGluValGlySerLeuAlaGluVal 471
Db 866 AACAGCTCGGTGAAGACGGGCAATGAGTCTGGCTCCAACTCTCGGACAGTGTCTGTG 925
QY 472 LysGluAsnProProPheTyr----- 478
Db 926 AAGCGTGTGACAGGACCTCCACCTGGGAGACCGTGTGCTGTGTGGGACCAAGACC 985
QY 479 GlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGlyLeuGlu 498
Db 986 GGTGTGTAGATATGTTGGGAGACGACTTGGCCAAAGCGGAGTGTGTGTGTGGAG 1045
QY 499 LeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyrPheThr 518
Db 1046 CTG---GACGAGCCCTTGGGAAGAACGATGGGCGAGTGGCGAGCGCAGGACTTCCAG 1102
QY 519 CysAlaLeuLysLysAlaLeuPheValLysLeu----- 529
Db 1103 TGCCCAACCAAGTTCGTCTCTTTGCCCAATCCCAAGGTCATCCGAATTTGGCTTCCCA 1162
QY 530 -----LysSerCysArgProAspSerArgPheAla----- 539
Db 1163 TCTACAGTCCAGCCCAAGGCCAAGAACCAAGCGCATGCTATGGGTGTCTCAGCCTTG 1222
QY 540 SerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAla-----PheGly 557
Db 1223 ACCCAGCCCGCAGCACTTCTCCATCAGCTCTGTGAGTCTGTGGCTCTCTGTGTGT 1282
QY 558 GlyTyrLeuSerGlu-----ValValGluGluAsnThrProProLysMetGluLys 574
Db 1283 GGCAGCGCCGAGCGTGTAGTGGCTGTCTCACAGAGACCTCTTACGCTATGCCCGAAGATC 1342
QY 575 GluGlyLeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyr-AsnSerCy 594
Db 1343 TCAGGCACACAGCCCTCGCAGGAGGCACCTGAAGGAGAAACAGCAGCAGCATCGAACAGCTG 1402
QY 594 sTyrLeuAspSerThr 599
Db 1403 CTGGCTGACGTGACT 1418
```

RESULT 8

```
US-60-660-557-55
; Sequence 55, Application US/60660557
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Ming-Yi Chiang
; APPLICANT: Nicholas M. Dean
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan M. Freier
; APPLICANT: Ravi Jain
; APPLICANT: Donna T. Ward
; TITLE OF INVENTION: COMPOSITIONS AND THEIR USES DIRECTED TO ORGANELLE COMPONENTS
; FILE REFERENCE: DPK-0075US.L
; CURRENT APPLICATION NUMBER: US/60/660,557
; CURRENT FILING DATE: 2005-03-10
; NUMBER OF SEQ ID NOS: 1853
; SOFTWARE: PatentSeq version 1.0
; SEQ ID NO 55
; LENGTH: 5793
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-660-557-55
Alignment Scores:
```

```
Pred. No.: 6.69 Length: 5793
Score: 151.00 Matches: 107
Percent Similarity: 31.58% Conservative: 61
Best Local Similarity: 20.11% Mismatches: 184
Query Match: 3.00% Indels: 180
DB: 12 Gaps: 23
```

US-09-671-687A-3 (1-949) x US-60-660-557-55 (1-5793)

```
QY 219 AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly 238
Db 300 GAGACTCAGGAGGAATTTGTGGATGACTTTTCGAGTGGGAGCGAGTTTGGTGATGCA 359
QY 239 -----GlyGluThrIleGluSerGlyThrValIlePhe 249
Db 360 AATAAGCCTGGATTTATCCAGTTTCTTGGAGAAACC----- 395
QY 250 CysAspValLeuProGlyLysGluSerLeuGlyTyrPheValGlyValAspMetAspAsn 269
Db 396 ---CAGTTTGCACAGGCCAG-----TGGGCTGGAAATTTGTTTAGATGAA 437
QY 270 ProIleGlyAsnTrpAspGlyArgPheAspGlyValLeuCysSerPheAlaCysValGlu 289
Db 438 CCATAGCAAGAACGATGTTCCGTGGCAGGAGTTCGG---TATTTCCAGTGTGAACCT 494
QY 290 SerThrIleLeuLeuHisIleAsnAspIleIleProGluSerValThrGlnGluArgArg 309
Db 495 -----TTAAAGGCGCATATTT-----ACCCGA 515
QY 310 ProProLysLeuAlaPheMetSerArgGlyValGlyAspLysGlySerSerHisAsn 329
Db 516 CCTTCAAAAGTTA-----ACAAGGAAGTG-----CAAGCA 545
QY 330 LysProLysAlaThrGlySerThrSerAspProGlyAsnArgArgSerGluLeuPheTyr 349
Db 546 GAAGATGAACCTAATGGCTGCAGACAAAGCCGCCCTCCCGAGCTACTTCCCGCTGTGC 605
QY 350 ThrLeuAsnGlySerSerValAspSerGlnProGlnSerLysSerLysAsnThrTrpTyr 369
Db 606 ACTTCTACGGCCAGCAGTGTGTCTTCTCCCTCCACCCCTTCAAAAC----- 653
QY 370 IleAspGluValAlaGluAspProAlaLysSerLeuThrGluIleSerThrAspPheAsp 389
Db 654 ATCCCTCAGAAACCATCACAGCCAGCA----- 680
QY 390 ArgSerSerProProLeuGlnProProValAsnSerLeuThrThrGluAsnArgPhe 409
Db 681 ---GCAAAGGAACCTTCAGCTACGCTCCGATCAGCAACCTTACA----- 722
QY 410 HisSerLeuProPheSerLeuThrLysMetProAsnThrAsnGlySerIleGlyHisSer 429
Db 722 ----- 722
QY 430 ProLeuSerLeuSerAlaGlnSerValMetGluGluLeuAsnThrAlaProValGlnGlu 449
Db 723 -----AAAACTGCCAGTGAATCTATCTCAACCTTTTCAGAGGCTGGCTCAATCAAGAAA 776
QY 450 SerProLeuAlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAla 469
Db 777 -----GGAGAAAGAGAGCTCAAAATCGGAGACAGATGA 809
QY 470 GluValLysGluAsnProProPheTyrGlyValIleArgTrpIleGlyGlnProGly 489
Db 810 TTGTTGGTGGCAGTAAAGCT-----GGTGTAGTCGGTTTCTTGGGAGACCGACTTT 863
QY 490 LeuAsnGluValLeuAlaGlyLeuGluLeuAspGluCysAlaGlyCysThrAspGly 509
Db 864 GCCAAGGGGAGTGGTGTGCGTGGAGTTA---GATGAGCCACTTGGGAGAAATGATGCG 920
QY 510 ThrPheArgGlyThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPheValLysLeu 529
Db 921 GCTGTTGTGGAACAAGGTATTTTTCAGTGTCAACCCAAATATGGCTTTGTTGCTCT----- 974
```

QY 530 LysSerCysArgProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIleGlu 549
DB 975 -----CCTGCCAC 983
QY 550 ArgCysAsnSerLeuAlaPheGlyGlyTyrLeuSerGluValValGluGlnAsnThrPro 569
DB 984 AAAGTTACCAAGATTGGCTTC-----CCTCCACTACACCA 1019
QY 570 ProLysMetGluLysGluGlyLeu---GluIleMetIleGlyLysLysLysGlyIleGln 588
DB 1020 GCCAAGCAAGCCCAAGCAGTGAGCGGAGTGATGGCAGCACGCTCCGCCAGCCTGAAG 1079
QY 589 GlyHisTyrAsnSerCysTyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSer 608
DB 1080 CGCAGCCCTTCTGCCTTCTCCCTCAGCTCC-----ATGAGCTCA 1118
QY 609 ValIleAspThrValLeuLeuArgProLysGluLysAsnAspValGluTyrTyrSerGlu 628
DB 1119 GTGGCTCTCTGTGAGCAGCAGCC-----AGTCGG 1151
QY 629 ThrGlnGluLeuLeuArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCys 648
DB 1152 ACAGGACTATTGACTGAACCTCTCCCGTTAGCCGAGGAGATCTCCGTACCACTGCC 1211
QY 649 AlaThrLysIleMetLys-----LeuArgLysIle 658
DB 1212 CTCAGGAGGCCCTGAAGGAGAGCAGCAGCATTGAGCAGCTGCTGGCGGAACGGAT 1271
QY 659 LeuGluLysValGluAlaAlaSerGlyPheThrSerGluLysAspProGluGluPhe 678
DB 1272 CTGAGAGGCGGAGGTGGCCAAAGGCC---ACAGGCCACGTGGGGAGATAGAGCAGGAG 1328
QY 679 LeuAsnIleLeu-----PheHisIleLeu 687
DB 1329 CTAGCTCTGGCCCGGAGCAGCAGATGACCAGCATGTCTGGAATGGAGCCAAATGGAC 1388
QY 688 -----ArgValGluProLeuLysIle 695
DB 1389 CAGCTCGCAACAATGTGGAGCTGCTGACAGGAGAGAGTGAGCTTCTCAACCAGCTT 1448
QY 696 ArgSerAlaGlyGlnLysValGlnAspCysTyrPhe 707
DB 1449 GAAGAGGAGAAAGGAAGGTGTGAGGACCTTCAGTTC 1484

RESULT 9

US-60-660-557-54
; Sequence 54, Application US/60660557
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Ming-Yi Chiang
; APPLICANT: Nicholas M. Dean
; APPLICANT: Kenneth W. Doble
; APPLICANT: Susan M. Freier
; APPLICANT: Ravi Jain
; APPLICANT: Donna T. Ward
; TITLE OF INVENTION: COMPOSITIONS AND THEIR USES DIRECTED TO ORGANELLE COMPONENTS
; FILE REFERENCE: DPK-0075US.L
; CURRENT APPLICATION NUMBER: US/60/660,557
; CURRENT FILING DATE: 2005-03-10
; NUMBER OF SEQ ID NOS: 1853
; SOFTWARE: PatentSeq version 1.0
; SEQ ID NO 54
; LENGTH: 5898
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-660-557-54

Alignment Scores:

Pred. No.:	6.84	Length:	5898
Score:	151.00	Matches:	107
Percent Similarity:	31.58%	Conservative:	61
Best Local Similarity:	20.11%	Mismatches:	184
Query Match:	3.00%	Indels:	180

DB: 12 Gaps: 23
US-09-671-687A-3 (1-949) x US-60-660-557-54 (1-5898)
QY 219 AspThrMetGlnValGluLeuProLeuGluIleAsnSerArgValSerLeuLysGly 238
DB 300 GAGACTCAGGAGGAATTTGGATGACTTTCGAGTTGGGAGCGAGTTTGGTGAATGGA 359
QY 239 -----GlyGluThrIleGluSerGlyThrValIlePhe 249
DB 360 AATAAGCCTCGATTATCCAGTTTCTTGGAGAAACC----- 395
QY 250 CysAspValLeuProGlyLysGluSerLeuGlyTyrPheValGlyValAspMetAspAsn 269
DB 396 ---CAGTTTGCACAGCCAG-----TGGGCTGGAATTTGTTAGATGAA 437
QY 270 ProIleGlyAsnTrpAspGlyArgPheAspGlyValLeuCysSerPheAlaCysValGlu 289
DB 438 CCATAGGCAAGAACGATGTTCCGTGGCAGGAGTTCGG---TATTTCCAGTGTGAACCT 494
QY 290 SerThrIleLeuLeuHisIleAsnAspIleIleProGluSerValThrGlnGluArg 309
DB 495 -----TTAAGGCGCATATTT-----ACCGA 515
QY 310 ProProLysLeuAlaPheMetSerArgGlyValGlyAspLysGlySerSerHisAsn 329
DB 516 CTTCAAAGTTA-----ACAAGGAGGTG-----CAAGCA 545
QY 330 LysProLysAlaThrGlySerThrSerAspProGlyAsnArgArgSerGluLeuPheTyr 349
DB 546 GAAGATGAAGCTAATGGCTGCAGACAAACCCGCTCCCGAGCTACTTCAACCGTGTGC 605
QY 350 ThrLeuAsnGlySerSerValAspSerGlnProLysSerLysSerLysAsnThrTyr 369
DB 606 ACTTCTACGCCAGCATGGTGTCTTCTCCCTCCCTCCACCTTCAAC----- 653
QY 370 IleAspGluValAlaGluAspProAlaLysSerLeuThrGluIleSerThrAspPheAsp 389
DB 654 ATCCCTCAGAAACCATCACAGCCAGCA----- 680
QY 390 ArgSerSerProLeuGlnProProValAsnSerLeuThrThrGluAsnArgPhe 409
DB 681 ---GCAAAAGGAACCTTCAGCTACGCTCCGATCAGCAACCTTACA----- 722
QY 410 HisSerLeuProPheSerLeuThrLysMetProAsnThrAsnGlySerIleHisSer 429
DB 722 ----- 722
QY 430 ProLeuSerLeuSerAlaGlnSerValMetGluGluLeuAsnThrAlaProValGlnGlu 449
DB 723 -----AAAACCTGCCAGTGAATCTATCTCCAACCTTTTCAGAGGTGGCTCAATCAAGAAA 776
QY 450 SerProLeuAlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAla 469
DB 777 -----GGAGAAAGAGAGCTCAAAATCGGAGACAGAGTA 809
QY 470 GluValLysGluAsnProProPheTyrValIleArgTrpIleGlyGlnProGly 489
DB 810 TTGGTTGGTGGCATAAGGCT-----GGTGTAGTCGGTTTCTTGGGAGACCGCACTTT 863
QY 490 LeuAsnGluValLeuAlaGlyLeuGluLeuAspGluCysAlaGlyCysThrAspGly 509
DB 864 GCCAAGGGGAGTGGTGGCGTGGAGTTA---GATGAGCCACTGGGAAGAATGATGCG 920
QY 510 ThrPheArgGlyThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPheValLysLeu 529
DB 921 GCTGTTGCTGGAAACAGGTATTTTCAGTGTCAACCCAAATATGGCTTGTTCGCT----- 974
QY 530 LysSerCysArgProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIleGlu 549
DB 975 -----CCTGCCAC 983
QY 550 ArgCysAsnSerLeuAlaPheGlyGlyTyrLeuSerGluValValGluAsnThrPro 569

Db 984 AAAGTTACCAAGATTGGCTTC-----CCTTCCACTACACCA 1019
Qy 570 ProLysMetGluLysGluGlyLeu---GluLeuMetIleGlyLysLysGlyIleGln 588
Db 1020 GCCAAAGCCCAAGCCCAAGCAGTGAGCGAGTATGTCGCGCAGCCTGAAG 1079
Qy 589 GlyHisTyrAsnSerCysTyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSer 608
Db 1080 CGCAGCCCTTCCTCCCTTCCTCCCTCAGCTCC-----ATGAGCTCA 1118
Qy 609 ValLeuAspThrValLeuLeuArgProLysGluLysAsnAspValGluTyrTyrSerGlu 628
Db 1119 GTGGCTCCTCTGTGAGCAGCAGGCC-----AGTCGG 1151
Qy 629 ThrGlnGluLeuLeuArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCys 648
Db 1152 ACAGGACTATTGACTGAACCTCTCCCGTTAGCCAGGAAGATCTCCGCTACCACTGCC 1211
Qy 649 AlaThrLysIleMetLys-----LeuArgLysIle 658
Db 1212 CTCACGAGCCCTGAAGGAGAGACAGCAGCAGCATTGAGCAGTGTGCGGGAACGGAT 1271
Qy 659 LeuGluLysValGluAlaLaserGlyPheThrSerGluGluLysAspProGluGluPhe 678
Db 1272 CTGGAGAGCGGAGGTGGCCAGGCC--ACAGCACCCTGGGGGAGATAGAGCAGGAG 1328
Qy 679 LeuAsnIleLeu-----PheHisIleLeu 687
Db 1329 CTAGCTCGCCCGGAGCAGCATGACAGCATGCTCTGGAATTGGAAGCCAAATGGAC 1388
Qy 688 -----ArgValGluProLeuLeuLysIle 695
Db 1389 CAGCTCGCAACATGTGGAGAGTGTGACAGGAGAGAGTGGAGCTTCTCAACCAAGCTT 1448
Qy 696 ArgSerAlaGlyGlnLysValGlnAspCysTyrPhe 707
Db 1449 GAAGAGGAGAAAGGAAGGTGGAGACCTTCAGTTTC 1484

RESULT 10
PCT-US05-07748-423
; Sequence 423, Application PC/TUS0507748
; GENERAL INFORMATION:
; APPLICANT: Avalon Pharmaceuticals
; TITLE OF INVENTION: Determining Cancer-Linked Genes and Therapeutic Targets Using
; TITLE OF INVENTION: Molecular Cytogenetic Methods
; FILE REFERENCE: 689290-237
; CURRENT APPLICATION NUMBER: PCT/US05/07748
; CURRENT FILING DATE: 2005-03-08
; PRIOR APPLICATION NUMBER: 60/550,304
; PRIOR FILING DATE: 2004-03-08
; NUMBER OF SEQ ID NOS: 3049
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 423
; LENGTH: 5449
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: cDNA Sequence
PCT-US05-07748-423

Alignment Scores:
Pred. No.: 13.7 Length: 5449
Score: 146.00 Matches: 112
Percent Similarity: 33.66% Conservative: 59
Best Local Similarity: 22.05% Mismatches: 189
Query Match: 2.90% Indels: 149
DB: 2 Gaps: 21
US-09-671-687A-3 (1-949) x PCT-US05-07748-423 (1-5449)
Qy 275 AspGlyArgPheAspGlyValLeuCysSerPheAlaCysValGluSerThr-IleLeuLe 294
|||||

270 GATGGCAGAGGAGAC-----GTGACCAGCACTCACCCCTTGT 305
Qy 294 uHisIleAsnAspIlelleProGluSerValThrGlnGluArgArgProProLysLeuAl 314
Db 306 CCACCTG-----CCAGTGGCACCCTATGAGAGAGCCAGCGCCTGAA 350
Qy 314 aPheMetSerArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaIth 334
Db 351 GCCCCCGCGCGTGGG-----GGGAGACCTCCACCCCATGGGCGG 392
Qy 334 rGlySerThrSerAspProGlyAsnArgSerGluLeuPheTyrThrLeuAsnGlySe 354
Db 393 GACATCTACTGGTTCAGCTTCATCTCGGGCGGCGTGGCGCTAGCTCCAGGAAGGCTC 452
Qy 354 rSerValAspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGluValAl 374
Db 453 CCCATGTCACAAACAGTCATCTGGACCTCTCTCTCCCG--GCCGAGCTGTGCTGCC 509
Qy 374 aGluAspProAlaLysSerLeuThrGluIleSerThrAspPhe----- 388
Db 510 CGAAGAGCCGGGCCCCAAGCGCGGAGAGTGGGGGATGACTTCTTGGGGGACTTTGTGGT 569
Qy 389 ---AspArgSerSerProProLeuGlnProProValAsnSerLeuThrThrGluAs 407
Db 570 GGGCGAGCGGTGTGGGTGAACCGCGTGAAGCCAGCGCTGTGTGAGTATCTGGGAGAC 629
Qy 407 nArgPhe----- 409
Db 630 GCAGTTGCGACCGCGGCGAGTGGGTGCTGCTGAGCAGCAGCGGTGGGCAAGATGA 689
Qy 410 -----HisSerLe 412
Db 690 TGGCGGCGTGGCGGCGTGCCTACTTTCGAGTGCCTCCGCGCTCCAGGCTATCTCACCG 749
Qy 412 uProPheSerLeuThrLysMetProAsnThr-----AsnGlySerIleGlyHisSerPr 430
Db 750 GCCTCTCAAGCTACCGCGGAGCCAGCGCGGAGGCTCGGGAGTGTATGCCACTCCTCGT 809
Qy 430 oLeuSerLeuSerAlaGlnSerValMetGluGluLeuAsnThrAlaProValGlnGluSe 450
Db 810 GGAGTGTGCTGACTGCCAGAACCTGTCTATTGCTTCGCGGACGCGC-----AC 857
Qy 450 rProProLeu-----AlaMetPr 456
Db 858 GCCCCCGCTACCGAGCGCGCTCATCCCTCGCGGAGAGCGCTCTCAACAGCTCCGTGAA 917
Qy 456 oProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPr 476
Db 918 GACTGGCAACAGAGTCGGGATCCAACTCTCAGACAGCGGCTCTGTGAAGCGGGCGGAAA 977
Qy 476 oPheTyr-----GlyValIleArgTr 483
Db 978 GGACCTGGCGCTGGGGAGCCGCGTGTGTTGGCGGAGCAGAGACTGGCGTGTGCGGTA 1037
Qy 483 pIleGlyGlnProProGlyLeuAsnGluValLeuAlaGlyLeuGluLeuGluAspGluCy 503
Db 1038 CGTGGGGAGACAGACTTTGCCAAGGCGAGTGTGTGGCGTGGAGCTG--GACGAGCC 1094
Qy 503 sAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyrPheThrCysAlaLeuLysLy 523
Db 1095 CTTGGGAAGAATGATGGGCGGCGTGGCGGCGCAGGAGTACTTCCAGTGCACCCCAAGTT 1154
Qy 523 sAlaLeuPheValLysLeu-----Ly 530
Db 1155 TGGTCTCTTCGCGCCCATCCAAAGTATCCGTATCGGCTTCCCATCTACAGCCAGC 1214
Qy 530 sSerCysArgProAspSerArgPheAla-----SerLeuGlnProVa 544
Db 1215 CAAGGCCAAGAAGACCAAGCGTATGGCCATGGGTGTGTGTCAGCACTGACCCACAGTCCCAG 1274
Qy 544 lSerAsnGlnIleGluArgCysAsnSerLeuAla-----PheGlyGlyTyrLeuSerGl 562
Db 1275 CAGTTTCCTCATCAGCTCCGTCCTGCTGTGGCTCTCTCGTGGGGGTGCGGCCAGCGC 1334

QY 562 u-----ValValGluGluAsnThrProProLysMetGluLysGluGlyLeuGluI 579
 Db 1335 CAGTGGCTGTCTCAGGAGACCTTCTACGCTACGCCCGCAAGATCTCGGACACACGCGC 1394
 QY 579 eMetileGlyLysLysGlyIleGlnGlyHis-TyrAsnSerCysTyrLeuAspSerT 599
 Db 1395 CTTCAGAGAGGACCTGAAGAGAGACGACGACATTTGAGGCTGTCTGGCTGAACGAGA 1454
 QY 599 hr-----LeuPheCysLeuPheAlaPheSerSerValLeuAspThr----- 612
 Db 1455 CCTGGAACGGGCTGAG-GTGGCCAAAGGCCACAAAGCCACATCTCGAGGTGGAGAGGAGA 1513
 QY 613 --ValLeuLeuArgProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGlu- 631
 Db 1514 TTGCCCTGTCTCAAGGCACAGCATGAG-----CAGTATGTTCAGAAAGCCGAGGAGA 1564
 QY 632 --LeuLeuArgThrGluIleValAlaPheProLeuArgIleTyrGlyTyrValCysAlaThrL 651
 Db 1565 AGCTGACGAGCGCCGGCTGCTC----- 1587
 QY 651 ysIleMetLysLeuArgLysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerG 671
 Db 1588 --GTGGAGAGCGTGGGAAA-----GAGAAGTGGACCTGTCCAAACACAGCTGGAGGAG 1639
 QY 671 luGluLysAspProGluGlu 677
 Db 1640 AGAGGAGGAAGGTGGAGGAT 1659

RESULT 11

PCT-US05-07748-422
 ; Sequence 422, Application PC/TUS0507748
 ; GENERAL INFORMATION:
 ; APPLICANT: Avalon Pharmaceuticals
 ; TITLE OF INVENTION: Determining Cancer-Linked Genes and Therapeutic Targets Using
 ; TITLE OF INVENTION: Molecular Cytogenetic Methods
 ; FILE REFERENCE: 689290-237
 ; CURRENT APPLICATION NUMBER: PCT/US05/07748
 ; CURRENT FILING DATE: 2005-03-08
 ; PRIOR APPLICATION NUMBER: 60/550,304
 ; PRIOR FILING DATE: 2004-03-08
 ; NUMBER OF SEQ ID NOS: 3049
 ; SOFTWARE: PatentIn version 3.3
 ; SEQ ID NO 422
 ; LENGTH: 5554
 ; TYPE: DNA
 ; ORGANISM: Artificial
 ; FEATURE:
 ; OTHER INFORMATION: cDNA Sequence
 PCT-US05-07748-422

Alignment Scores:
 Pred. No.: 14 Length: 5554
 Score: 146.00 Matches: 112
 Percent Similarity: 33.66% Conservative: 59
 Best Local Similarity: 22.05% Mismatches: 189
 Query Match: 2.90% Indels: 149
 DB: 2 Gaps: 21

US-09-671-687A-3 (1-949) x PCT-US05-07748-422 (1-5554)

QY 275 AspGlyArgPheAspGlyValLeuCysSerPheAlaCysValGluSerThr-IleLeuLe 294
 Db 270 GATGGCAGAGAGAC-----GTGACACGACCTACCCCTTGT 305
 QY 294 uHisIleAsnAspIleIleProGluSerValThrGlnGluArgArgProProLysLeuAl 314
 Db 306 CCACCTG-----CCCAGTGGCACCCGATGCAGAAAGCCCGGCGCTGAA 350
 QY 314 aPheMetSerArgGlyValGlyAspLysGlySerSerSerHisAsnLysProLysAlaTh 334
 Db 351 GCGCCCGCGCGGTGGG-----GGGAAGCACTCCAGCCCGCATGGGCCG 392

QY 334 rGlySerThrSerAspProGlyAsnArgArgSerGluLeuPheTyrThrLeuAsnGlySe 354
 Db 393 GACATCTACTGGTTCAGCTTCTTCGCGCGGTGGCGCTAGCTTCCAAGGAAGCTC 452
 QY 354 rSerValAspSerGlnProGlnSerLysSerLysAsnThrTyrTyrIleAspGluValAl 374
 Db 453 CCCACTGCAACAACAGTCATCTGGACCTCTCTCTCTCCCG--GCCGACGCTGCTGCC 509
 QY 374 aGluAspProAlaLysSerLeuThrGluIleSerThrAspPhe----- 388
 Db 510 CGAGAAGCCGGGCCCCAAGCGCGAGAGTGGGGGATGACTTCTCGGGGACTTTGTGGT 569
 QY 389 ----AspArgSerSerProProLeuGlnProProProValAsnSerLeuThrThrGluAs 407
 Db 570 GGGCGAGCGGTGTGGGTGAACGGCGTGAAGCCAGCGGTGGTGCAGTATCTGGGAGAGAC 629
 QY 407 nArgPhe----- 409
 Db 630 GCAGTTTCGACCGGCGCAGTGGCTGGCTGGTGTGGTGTGGACGCCGCTGGGCAAGAATGA 689
 QY 410 -----HisSerLe 412
 Db 690 TGGCGGTGGGCGGTGGCTTCTGAGTCCCGGCCCTCCAGGGTATCTTCACGCG 749
 QY 412 uProPheSerLeuThrLysMetProAsnThr-----AsnGlySerIleGlyHisSerPr 430
 Db 750 GCCCTCAAAGCTGACCCCGCAGCCAGCGCGAGGGTCTCGGGAGTGTATGCCACTCCGT 809
 QY 430 oLeuSerLeuSerAlaGlnSerValMetGluGluLeuAsnThrAlaProValGlnGluSe 450
 Db 810 GGAGTGCCTGACTGCCCAAGAACCTGTCATTCGATTCGGGCACGCGC-----AC 857
 QY 450 rProProLeu-----AlaMetPr 456
 Db 858 GCGCCCGCTGACCGCGCGCTCATCCCTCCGCGGAGAGCGTCTCTCAACAGCTCCGTA 917
 QY 456 oProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPr 476
 Db 918 GACTGGCAACAGTGGGATCCAACTCTCAGACAGCGGCTCTGTGAAGCGGGCGGCAAAA 977
 QY 476 oPheTyr-----GlyValIleArgTr 483
 Db 978 GGACCTGGCGCTGGGGGACCGCGTGTGGTGGCGGAGCAAGACTGCGGTGGTGGCGTA 1037
 QY 483 pIleGlyGlnProProGlyLeuAsnGluValLeuAlaGlyLeuGluLeuGluAspGlyCu 503
 Db 1038 CGTGGGGGAGACAGACTTTCCCAAGCGCGAGTGGTGGCGTGGAGCTG--GACGAGCC 1094
 QY 503 sAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyrPheThrCysAlaLeuLysLy 523
 Db 1095 CCTTGGGAAGATGATGGGCGGTGGCGGCGACAGGTACTTCCAGTCCGCCACCAAGTT 1154
 QY 523 sAlaLeuPheValLysLeu-----Ly 530
 Db 1155 TGGTCTCTTCGCGCCCATCCAAAGTATCGGTATCGGTTCCTCATCTACCCAGCCGAGC 1214
 QY 530 sSerCysArgProAspSerArgPheAla-----SerLeuGlnProVa 544
 Db 1215 CAGGCCAAGAGACCCAGCGGTATGGCCATGGGTGTGTGCAGCTGACCCACAGCTCCAG 1274
 QY 544 lSerAsnGlnIleGluArgCysAsnSerLeuAla-----PheGlyGlyTyrLeuSerGl 562
 Db 1275 CAGTCTCTCAGCTCCGCTCAGCTCTGTGGCTCTCTCGTGGGGGTCTCGGCCAGCGC 1334
 QY 562 u-----ValValGluGluAsnThrProProLysMetGluLysGluGlyLeuGluI 579
 Db 1335 CAGTGGCTGTCTCAGGAGACCTTCTTACGCTACGCCCGCAAGATCTCGGCAACACGCGC 1394
 QY 579 eMetIleGlyLysLysGlyIleGlnGlyHis-TyrAsnSerCysTyrLeuAspSerT 599
 Db 1395 CTTCAGAGGAGCACTGAAGAGAGAGCAGCAGCATTTGAGCAGCTGTCTGGCTGAACGAGA 1454
 QY 599 hr-----LeuPheCysLeuPheAlaPheSerSerValLeuAspThr----- 612

```

Db      1455  CCTGGAACGGCTGAG-GTGGCCAGGCCACAGCCACATCTCGGAGTGGAGAGGAGA 1513
Qy      613  --ValLeuLeuArgProLysGluYsAsnAspValGluTyrTyrSerGluThrGlnGlu- 631
Db      1514  TTCCCTGCTCAAGGACACAGCATGAG-----CAGTATGTTGCAGAACGCCGAGGAGA 1564
Qy      632  --LeuLeuArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrL 651
Db      1565  AGCTGCAGCGAGCCCGCTGCTC----- 1587
Qy      651  ysIleMetLysLeuArgLysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerG 671
Db      1588  --GTGGAGAGCGTGGGAAA-----GAGAGGTGGACCTGTCCAACACAGCTGGAGGAGG 1639
Qy      671  luGluLysAspProGluGlu 677
Db      1640  AGAGGAGGAAGTGGAGGAT 1659

```

RESULT 12

```

PCT-US05-07748-705
; Sequence 705, Application PC/TUS0507748
; GENERAL INFORMATION:
; APPLICANT: Avalon Pharmaceuticals
; TITLE OF INVENTION: Determining Cancer-Linked Genes and Therapeutic Targets Using
; FILE OF INVENTION: Molecular Cytogenetic Methods
; FILE REFERENCE: 689290-237
; CURRENT APPLICATION NUMBER: PCT/US05/07748
; PRIOR FILING DATE: 2005-03-08
; PRIOR APPLICATION NUMBER: 60/550,304
; PRIOR FILING DATE: 2004-03-08
; NUMBER OF SEQ ID NOS: 3049
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 705
; LENGTH: 5555
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: cDNA Sequence
PCT-US05-07748-705

```

```

Alignment Scores:
Pred. No.:      14      Length:      5555
Score:          146.00  Matches:      112
Percent Similarity: 33.66%  Conservative: 59
Best Local Similarity: 22.05%  Mismatches: 189
Query Match:      2.90%  Indels:      149
DB:                2      Gaps:        21

```

US-09-671-687A-3 (1-949) x PCT-US05-07748-705 (1-5555)

```

Qy      275  AspGlyArgPheAspGlyValLeuCysSerPheAlaCysValGluSerThr-IleLeuLe 294
Db      270  GATGGCAGAGGAGAC-----GTGACCAGCATCACCCTTGT 305
Qy      294  uHisIleAsnAspIleileProGluSerValThrGlnGluArgProLysLeuAl 314
Db      306  CCACCTG-----CCCATGCGACCGCCATGCAGAACGCCAGCGCCTGAA 350
Qy      314  aPheMetSerArgGlyValGlyAspLysGlySerSerSerHisAsnLysProLysAlaTh 334
Db      351  GCCCCCGCGCGTGGG-----GGGAAGCACTCCAGCCCCATGGGCCG 392
Qy      334  rGlySerThrSerAspProGlyAsnArgArgSerGluLeuPheTyrThrLeuAnGlySe 354
Db      393  GACATCTACTGGGTCACTTCCTCTCGCGCGGGTGGCCGCTAGCTCCCAAGGAAGGCTC 452
Qy      354  rSerValAspSerGlnProGlnSerLysSerLysAsnThrTyrTyrIleAspGluValAl 374
Db      453  CCACCTGCACAACAGTCATCTGGACCTCTCTCTCTCCCG---GCCCGACGTGCTGCCCC 509
Qy      374  aGluAspProAlaLysSerLeuThrGluIleSerThrAspPhe----- 388

```

```

Db      510  CGAAGCCGCGGCCCAAGCGCGGGAAGTGGGGATGACTTCTCTGGGGGACTTTGTGTT 569
Qy      389  ----AspArgSerSerProLeuGlnProProValAsnSerLeuThrThrGluAs 407
Db      570  GGGCGAGCGGGTGTGGGTGAACCGCGTGAAGCCAGCGGTGTGTGAGTATCTGGGAGAGAC 629
Qy      407  nArgPhe----- 409
Db      630  GCAGTTCGCACCGGGCCAGTGGGGTGGGTGCTGGACGACCCGGTGGGCAAGAATGA 689
Qy      410  -----HisSerLe 412
Db      690  TGGCGGGTGGCGGGCGTGCCTACTTTCGAGTCCCGGCCCTCCAGGGGTATCTTCCACGCG 749
Qy      412  uProPheSerLeuThrLysMetProAsnThr-----AsnGlySerIleGlyHisSerPr 430
Db      750  GCCTCTCAAGCTGACCCGCGCAGCGCTCGGGAGGTGATGCCCATCTCCCT 809
Qy      430  oLeuSerLeuSerAlaGlnSerValMetGluGluLeuAsnThrAlaProValGlnGluSe 450
Db      810  GGAGTGCCTGCTACTGCCCGAAGAACTGTCTATTGCTTCGGGGCACGGCC-----AC 857
Qy      450  rProProLeu-----AlaMetPr 456
Db      858  GCCCGCGCTACACGCGCGCTCATCCCTCGCGGAGAGCGCTCTCAACAGCTCCGTGAA 917
Qy      456  oProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPr 476
Db      918  GACTGGCAACGAGTGGGATCCAACTCTCAGACAGCGGCTCTGTGAAGCGGGCGGCAAAA 977
Qy      476  oPheTyr-----GlyValIleArgTr 483
Db      978  GGACCTGCGCTCGGGGACCGCGTGTGGTGGCGGAGCGAAGACTGCGGTGGTGGCGTA 1037
Qy      483  pIleGlyGlnProProGlyLeuAsnGluValLeuAlaGlyLeuGluLeuGluAspGluCy 503
Db      1038  CTGGGGGAGACAGACTTTGCCAAGGGCGAGTGTGGGTGGGTGAGCTG---GACGAGCC 1094
Qy      503  sAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyrPheThrCysAlaLeuLysLy 523
Db      1095  CCTTGGGAAGAATGATGGGGCGGTGGCGGCGCACAGGTACTTTCAGTGCACCACCAAGTT 1154
Qy      523  sAlaLeuPheValLysLeu-----Ly 530
Db      1155  TGGTCTCTTCGCGCCCATCCACAAGTGATCCGTATCGGCTTCCCATCTACCAGCCAGC 1214
Qy      530  sSerCysArgProAspSerArgPheAla-----SerLeuGlnProVa 544
Db      1215  CAAGGCCAAGAACACCAAGCGTATGGCCATGGGTGTGTACGACTGACCCACAGTCCCAG 1274
Qy      544  lSerAsnGlnIleGluArgCysAsnSerLeuAla-----PheGlyGlyTyrLeuSerGl 562
Db      1275  CAGTTCCTCATCAGCTCCGTCGCTCTGTGGCCTCTCCGTGGGGGTTCGCCAGCGCG 1334
Qy      562  u-----ValValGluGluAsnThrProProLysMetGluLysGluGluI 579
Db      1335  CAGTGGCGCTCTCAGCGAGACCTTTCAGCTACGCCCGCAAGATCTCGGGCACCCAGCGC 1394
Qy      579  eMetIleGlyLysLysGlyIleGlnGlyHis-TyrAsnSerCysTyrLeuAspSerT 599
Db      1395  CTTGCAGGAGGCACTGAAGGAGAGAGCAGACATTCAGCAGCTGCTGGGTGACGAGA 1454
Qy      599  hr-----LeuPheCysLeuPheAlaPheSerSerValLeuAspThr----- 612
Db      1455  CTGGGAACGGGCTGAG-GTGGCCNAGGCCACAGCCACATCTGCGAGTGGAGAGGAGA 1513
Qy      613  --ValLeuLeuArgProGlyLysAsnAspValGluTyrTyrSerGluThrGlnGlu- 631
Db      1514  TTGCCCTGCTCAAGGACACAGCATGAG-----CAGTATGTTGCAGAACGCCGAGGAGA 1564
Qy      632  --LeuLeuArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrL 651
Db      1565  AGCTGCAGCGAGCCCGCTGCTC----- 1587

```



```
-QY 651 yalleMetLysLeuArgLysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerG 671
Db 1588 --GTGGAGACGTCGCGGAA--GAGAAGGTGGACCTGTCCAACCCAGCTGGAGGAGG 1639
QY 671 luGluLysAspProGluGlu 677
Db 1640 AGAGGAGGAAGTGGAGGAT 1659

RESULT 13
US-10-932-182A-77607
; Sequence 77607, Application US/109322182A
; GENERAL INFORMATION:
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHIISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: 030685-043
; CURRENT APPLICATION NUMBER: US/10/932,182A
; CURRENT FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 77607
; LENGTH: 2154
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-10-932-182A-77607

Alignment Scores:
Pred. No.: 21 Length: 2154
Score: 136.00 Matches: 82
Percent Similarity: 39.38% Conservative: 46
Best Local Similarity: 25.23% Mismatches: 126
Query Match: 2.70% Indels: 72
DB: 7 Gaps: 15

US-09-671-687A-3 (1-949) x US-10-932-182A-77607 (1-2154)

QY 392 SerProProLeuGlnProProProValAsn-----SerLeuThrThrGluAsnArgPhe 409
Db 412 TCCTTACCGCTACGGCTTTTACCTGTTCATATGATCGCTCTTGTCAATACATAACAAGTTT 471
QY 410 -----HisSerLeuProPheSerLeuThrLysMetProAsnThrAsnGlySer 425
Db 472 GATGGCTCACTACATGAATACCAATGACATGACCTAACAAACCAACAATGATAAT----- 525
QY 426 IleGlyHisSerProLeuSerLeuSerAlaGlnSerValMetGluGluLeuAsnThrAla 445
Db 526 -----TCGAAAGAGGACATAGTCAGGGAAGTAATCAGATA 561
QY 446 ProValGlnGluSerProProLeuAlaMetProGlyAsnSerHisGlyLeuGluVal 465
Db 562 -----GCATCAAGTAATAAATTCGAAGCT 585
QY 466 GlySerLeuAlaGluValLysGluAsnProPhePheGlyValIleArgTyrPheGly 485
Db 586 GGTTC-----GAAGTG-----GCCTACTACACACAGCAAAAGGATTATCA 627
QY 486 GlnProGlyLeu-AsnGluValLeuAlaGlyLeuGluLeuGluAspGluCys----- 503
Db 628 AAACCTTCGTATTAAACTCAGTACTGGAAGAGTCAATATTCAAACGCTGTCTAGT 687
QY 504 -----AlaGlyCysThrAspGlyThrPheArgGlyThrArgTyrPheThrCys 519
Db 688 CCAGCAACTGCTCCCCCGTCACTCTTGGAGGTTTCGTCAACAATCAGAGATTCTAGC 747
QY 519 sAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSerArgPheAl 539
Db 748 CAAGATTCTCATCTTCTATCCAGGTGGAACCAAGAGGAG-GAA----- 795
QY 539 aSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPheGlyGlyTy 559
```

```
Db 796 -----GAAGGAAGATCGAGGCCATAGATCTTCGCCCCCGAAGCTTA 839
QY 559 rLeuSerGluValGluGluAsnThrProProLysMetGluLysGluGlyLeuGluI 579
Db 840 CAATCTACCAGTAATTGAAGATCTTAATGATCTGTTATCAGAG-----TTATCTAT 890
QY 579 eMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeuAspSerTh 599
Db 891 T-----ACTGGCTGCAAAATCCATGCAATACATGCTATATTAATAGTAT 935
QY 599 rLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArgProLysG 619
Db 936 AATTCAATGCTCTTTGGCACTACCTTATTTCGGGATTTCTTTTAAACCAAGAAATATAG 995
QY 619 u-----LysAsnAspValGluTyrSerGluThrGlnGluLeuLeuArgThrGluI 637
Db 996 GCTATTCTCTCAATACTAACAATAATCCGAAGAAGTCCAA--CTTCCCGCTCGATT 1052
QY 637 eValAsnProLeuArgIleTyr-----GlyTyrValCysAlaThrLysIleMe 653
Db 1053 CGTTTGTTCAAAAAATGATTGTAATGGAGGAGGAGGATTATACCAATAGATTCTT 1112
QY 653 tLysLeuArgLysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGlu 673
Db 1113 AAAAATGTGCAAG-----AAGTTAAGGCCAGATTTGAACATCCAGATGATCAACA 1163
QY 673 sAspProGluGluPheLeuAsnIleLeu-----PheHisIleLeuArgValG 690
Db 1164 GGATACAGGAGTTTCTACTAATAGTTTGTAGCAGCAATACATGAGAGCTCTCGAACA 1223
QY 690 uProLeuLeuLys 694
Db 1224 AATGTTGTCAAG 1236

RESULT 14
US-10-450-763-21942
; Sequence 21942, Application US/10450763
; GENERAL INFORMATION:
; APPLICANT: HVEseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 21942
; LENGTH: 1497
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (1156)..(995)
; OTHER INFORMATION: 38% homologous to Homo sapiens A secreted protein encoded by
; OTHER INFORMATION: clone dt674_2, accession number W80408, Smith-Waterman Score=87.
US-10-450-763-21942

Alignment Scores:
Pred. No.: 47.4 Length: 1497
Score: 128.00 Matches: 111
Percent Similarity: 34.89% Conservative: 68
Best Local Similarity: 21.64% Mismatches: 183
Query Match: 2.54% Indels: 152
DB: 8 Gaps: 24

US-09-671-687A-3 (1-949) x US-10-450-763-21942 (1-1497)
```

QY 218 GlyAspThrMetGlnValGluLeuProProLeuGluIle-----AsnSerArgVal 234
DB 20 GGGGAAACATGCGAGTCTCGCTTCTTCTTGACCATATGCTGTGCTTCTCGGGT 79
QY 235 SerLeuLysGlyGlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuPro 254
DB 80 TCAGTGAATCAGGCTTCACT-----GGNAGTGTGCCATCGCTGTAGACTCTCC 130
QY 255 GlyLysGluSerLeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrp 274
DB 131 GGACTCAATTCCTC-----TCTCCCTCCCGTGATAATCTAATCGAATTTT 178
QY 275 AspGlyArgPheAspGlyValLeuCysSerPheAlaCysValGluSerThrIleLeuLeu 294
DB 179 TCAGCT-----GCTCGTCTCTCGCTGCCATCCAGCCCA-----CTC 214
QY 295 HisIleAsnAspIleProGluSerValThrGlnGluArgArgProProLysLeuAla 314
DB 215 CCAGTGGCTGCTCTCGGCCCTTGACCTTCTCTTGCTCTGCTAGCCCTGAGGTC 274
QY 315 PheMetSerArgGlyValGly-----TTCGGCTCCAGGCTCTCGAGTGGCCACCCAGCCTCTGTTGTATATGTCACCTGTCATC 334
DB 275 TCCGGCTCCAGGCTCTCGAGTGGCCACCCAGCCTCTGTTGTATATGTCACCTGTCATC 334
QY 322 -----AspLysGlySerSerSerHisAsnLysProLysAlaThrGlySerThr----- 337
DB 335 TCCTCTGCTGCTGGCACTTCTCCCACTCCCGTGCTAATTCCTCTACCACTCCCTCCCT 394
QY 338 -----SerAspProGly 341
DB 395 GGAGCTGCTGTTTTCTTCTACTACTCATCTCTGTTGTGGTGGCAGGTTTCTCTCCAGCAACA 454
QY 342 AsnArgArgSerGluLeuPheThrThrLeuAsnGlySerSerValAspSerGlnProGln 361
DB 455 TCTGAAAGAGGAGGCTTATTGTGCTATGTTTCTCTCGCTCTCTCGAAGCTGCC 514
QY 362 SerLysSerLysAsnThrTrpTrpIleAspGluValAlaGluAspProLysSerLeu 381
DB 515 AACACCTCCCTCTCTCTCGGCTGTTTCTCGCGTTTCCACCACTCCGTGTAGAACTCC 574
QY 382 ThrGluLysSerThrAsp-----PheAspArgSerSer 392
DB 575 ACAGCCACAACTCTCTCTCCATGCCCTCATCTTTTGTGTGTT-----CCTCT 628
QY 393 ProProLeuGlnProPro-----ValAsnSerLeuThrThrGluAsnArgPhe 409
DB 629 CTTCCATGTCCTCGCTCCCTCCGAGTCTCTGCTGCTGCTCTCTCTCCAGAGT 688
QY 410 HisSerLeuProPheSerLeuThrLysMetProAsnThrAsnGlySerIleGlyHisSer 429
DB 689 CACTCTTTGCTCTCTCTCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 748
QY 430 ProLeu-----SerLeuSerAlaGlnSerValMetGluLeuLeuAsnThr--- 444
DB 749 GCCATTATGGGATCTTCTCTCAGCAACAGCTTCTCTGGGTCTCTAGCAGCATCCACTCCT 808
QY 445 -----AlaProValGlnGluSerProProLeuAla-MetPr 456
DB 809 CTTGGCTTCACT 868
QY 456 oProGlyAsnSerHisGlyLeuGluValGlySer----- 467
DB 869 TCTT-----CTGCTGGCTGCTGGCTGGGATCAGCCCTCAGGGGCTGACCTTTCTCT 922
QY 468 -LeuAlaGluValLysGlu-----AsnProProPheTyr----- 478
DB 923 GCTGGCTCAGCTTCAAGAGCTGTAAGGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 982
QY 479 -----GlyValIleArgTrpIle 484
DB 983 AGGGCCCTATTCTCTCTCTGTCAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1042
QY 484 eGlyGlnProProGlyLeu-----AsnGluValLeuAlaGlyLeuGluGln 500

DB 1043 TCACTCAGTCTTTCCCT 1102
QY 500 uAspGluCysAlaGlyCysThrAspGlyThr-----PheArgGlyTh 514
DB 1103 AATACCAGGCGTGTCTCAATATATGGCAACCACTGCTCTCTTAGTGATCTTCTCTGGTTG 1162
QY 514 rArgTyrPheThrCysAlaLeuLysLysAlaLeuPheValLysLysSerCysArgPr 534
DB 1163 GAGGTATGTCACCTGCTAGGAGATCAAGACA----- 1193
QY 534 oAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCys---AsnSe 553
DB 1194 ----ACAGCATGGCTGAAAGAACCTTCAAGAGGTCAAAACATCGCTGTAGCATGG 1249
QY 553 rLeuAlaPheGlyGlyTyrLeuSerGluValValGlu-----GluAsnTh 568
DB 1250 TGTCTCGAAGCTGGAAGGTTCTCACAGCTGCAGATGGTCAATAAGATGACACAGT 1309
QY 568 rProProLysMetGluLysGluGluGluIleMetIleGlyLysLysGlyIleGln 588
DB 1310 TGGCCCAAGAGTAATCGT---GGCCATAGTGTCTGTTGAT-----GA 1351
QY 588 nGlyHisTyrAsnSerCysTyrLeu-----AspSerThrLeuPheCysLe 603
DB 1352 AGGACATTCAGTCCGTTAGCACTCTAAACCTCTCTCAGAGATGGGACCATCTCTTCTCTCT 1411
QY 603 uPheAlaPheSerSerValLeuAspThrValLeuLeu 615
DB 1412 GCTTGCTTG-AAATTCCTTTTCAACTCTACTTTTA 1447

RESULT 15

US-10-467-657-5405/c
; Sequence 5405, Application US/10467657
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 5405
; LENGTH: 3918
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-5405

Alignment Scores:
Pred. No.: 183 Length: 3918
Score: 127.00 Matches: 173
Percent Similarity: 30.98% Conservative: 86
Best Local Similarity: 20.69% Mismatches: 267
Query Match: 2.52% Indels: 312
DB: 7 Gaps: 37

US-09-671-687A-3 (1-949) x US-10-467-657-5405 (1-3918)

QY 121 GlyLeuGlnIleAspValGlyCysProVal-----LysValGlnLeu 134
DB 1974 GGGCGCGCGTGGAC---GGCTGTACCGCGTGATTTGTAGTCGGCAAGCGCGCATC 1918
QY 135 ArgSerGlyGluLysPheProGlyValValArgPheArgGlyProLeuLeuAlaGlu 154
DB 1917 CAAATGGCGAGCAGCAGTCCCA-----ACGCGCTTTGACGCGGCAACGCTGGCGAG 1864
QY 155 ArgThrValSerGlyIlePhePheGlyValGluLeuGluGlyArg-----Gly 172


```

Qy 787 -----LeuAlaMetTyrGluCysArgGluCysTyrAspAspProAspIleSerAlaGly-- 804
Db 152 TCGGTATAAAGATATAGTTGTAGGGAATTT-----CCGAAATACGCTCGGGGCG 102
Qy 805 -----LysIleLysGlnPheCysLysThrCysAsnThrGlnVa 817
Db 101 GTAGTCGTGTCATGCTGTGTCGGTGTTCGGCAGTTTTTTATTGAAATGC----- 53
Qy 817 lHisLeuHisProLysArgLeuAsnHisLysTyrAsnPro 830
Db 52 -----CGTTTGAAAGTCGGTTTCAGCCG 29

```

Search completed: April 18, 2005, 22:23:19
Job time : 650 secs

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	151	3.0	1427	4	US-09-538-092-1044		Sequence 1044, Ap
2	135	2.7	1043	4	US-09-949-016-11650		Sequence 11650, A
3	130.5	2.6	484	4	US-09-248-796A-24175		Sequence 24175, A
4	125.5	2.5	1469	4	US-09-262-537-58		Sequence 58, Appl
5	118.5	2.4	1466	4	US-09-262-537-20		Sequence 20, Appl
6	118.5	2.4	1471	4	US-08-811-519-1		Sequence 1, Appl
7	118	2.3	267	3	US-09-399-913-57		Sequence 57, Appl
8	118	2.3	267	3	US-09-350-614-57		Sequence 57, Appl
9	118	2.3	677	4	US-09-270-767-45699		Sequence 45699, A
10	117	2.3	386	4	US-09-248-796A-17565		Sequence 17565, A
11	117	2.3	798	2	US-08-222-617A-8		Sequence 8, Appl
12	115.5	2.3	719	4	US-09-949-016-7766		Sequence 7766, Ap
13	114	2.3	523	4	US-09-248-796A-17555		Sequence 17555, A
14	114	2.3	850	4	US-09-129-603-2		Sequence 2, Appl
15	114	2.3	2777	4	US-10-220-587-4		Sequence 4, Appl
16	113	2.2	1226	1	US-08-280-443-2		Sequence 2, Appl
17	113	2.2	1226	1	US-08-457-453-2		Sequence 2, Appl
18	113	2.2	1226	1	US-08-555-678-2		Sequence 2, Appl
19	113	2.2	1226	5	PCI-US95-02275-2		Sequence 2, Appl
20	112.5	2.2	493	3	US-08-999-774A-12		Sequence 12, Appl
21	112.5	2.2	805	4	US-09-538-092-257		Sequence 257, App
22	112	2.2	734	4	US-09-328-352-4412		Sequence 4412, Ap
23	112	2.2	869	4	US-09-902-540-10125		Sequence 10125, A
24	111.5	2.2	1317	4	US-09-949-016-7588		Sequence 7588, Ap
25	110.5	2.2	657	3	US-09-370-368-7		Sequence 7, Appl
26	110	2.2	577	4	US-09-949-016-11572		Sequence 11572, A
27	109.5	2.2	868	4	US-09-538-092-787		Sequence 787, App

QY 560 LSEVVEENTPPKMEKGL-EIMIGKKKGIOGHYNHSCVLDSTLCLFAPSSVLDTVLLRPK 618
Db 284 -----ESTTPAKAKANAVRRVMAATTSASLKRSPSASSLS-----MSSVASSVSRP- 330
QY 619 EKNDVEYSTQELLTEIWNPLRIYGVYCATIMK-----LRKILEVEAASGF 668
Db 331 -----SRTGLTSTSSRYARKISGTTALQALKEKQOHIQOLLAERDLERAFAKA- 381
QY 669 TSEKPEEPFNIL-----FHHIL-----RVEPLLKIRSAQKQVQDC 705
Db 382 TSHVGEIQELALARDGHDQHVLEBAKMDQLRTWVEADREKVELLNQLEBEKRVEDL 441
QY 706 YF 707
Db 442 QF 443

RESULT 2

US-09-949-016-11650
; Sequence 11650, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ FOR Windows Version 4.0
; SEQ ID NO 11650
; LENGTH: 1043
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11650

Query Match 2.7%; Score 135; DB 4; Length 1043;
Best Local Similarity 22.6%; Pred. No. 0.0051;
Matches 66; Conservative 26; Mismatches 112; Indels 88; Gaps 10;
QY 309 RPKLAFMSRGVGDGKSSSHNKPATGSTDGPNRRSELFTYTLNGSSVDSQPSKSNW 368
Db 8 KPP-----GRGK-----HSSPMGRSTSGASSAAVAASKEGSPHKQSGSPSSP- 55
QY 369 YIDEVAEDPAKSLTEISTDF-----DRSSPPLQPPVNSLTENRF----- 409
Db 56 AAAAAPEKPGKAAEVGDDFLGVGERVYVNGVKRGVYQVYLGETOQAFQWAGVVLDD 115
QY 410 -----HSLPFLTKMPT-----NGSIGHSPLSLSAQSVMEELNT 444
Db 116 PVGNKDGAVGVRYFPCPALQGIPTRESKLTQPTAEGSGSDAHSVESLTQNLSHSGT 175
QY 445 APVQESPL-----AMPPGNHGLEVGSLAEVKENPPFY----- 478
Db 176 A-----TPPLTSRVILPRESVLNNGSVKTNESGNSLSDSGSVKRGKDLRLGDRVLVGGTK 231
QY 479 -GVIRWIGQPPGLNEVLAGELEDEACAGCTDGTFRGTRYFTCAKALFKVL 529
Db 232 TGVVRYVGETDFAKGECGVLEL-DEPLGKNDGAVAGTRYFCPPKFLGFAP 282

RESULT 3

US-09-248-796A-24175
; Sequence 24175, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:

; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 24175
; LENGTH: 484
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-24175

Query Match 2.6%; Score 130.5; DB 4; Length 484;
Best Local Similarity 22.2%; Pred. No. 0.0035;
Matches 128; Conservative 68; Mismatches 187; Indels 193; Gaps 35;
QY 317 SRGVGDKGSSHNKPKATG-----STSDPGNRRSELFTYTLNGSS-----VDSQPSKS 364
Db 39 SSNLDNSGTT-----KTSGHEIVNISKVPNGNEHSE--YEVNGVGLSILEHSSPFPKR 92
QY 365 KNTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSLTENRFHSLP--FSLTKMPT 422
Db 93 KTTYQERVSYIPNHDFTGKSSIFESSP--QP-----SIPNTFYKSHPNA 137
QY 423 NG-SIGHSPLSLSAQSVMEELNTAPVQESPLA--MPPGNHGLEVGSLAEVKENPPFYG 479
Db 138 KAKALGATQITEETSQEQEQN-GDVQ--PPLSTKQTPKNN-----KENT----- 180
QY 480 VIRMIGOPGLNEVLAGELEDEACAGCTDGTFRGTRYFTCAKALFKVLSKCRPDSRFA 539
Db 181 TPRWM--PPVLDERKNVQDLIED-----IKFT-----PSMPSIQQHH-- 215
QY 540 SLPQVSNQ-----IERCN--SLAFGGY--LSEVVEENTPPKMEK 574
Db 216 -QQPLSNQSRGGISPSPFQDFEHTGDTMHNNSVQSVETPGYRRARQDYKQKQPDMLK 274
QY 575 EGLEIMIGKKGIQG-----HYNICYLDSTLCLFAPSSVLDTVLLRPKE-KND 622
Db 275 SIPTVGGSGSGGGAGPDETNTKDHSTS-----STI-----SSMFTGIQKRPDEIKQQ 324
QY 623 VEYVSETOELLRTIYNPLRIYG--YVCATKIMKRLKILEKVEAASGFTSEKDPPEFLN 680
Db 325 METINEMKNMKNPE--SPLKLYGPKYNTYTR-NQLAGLVNLSKNKNTPAQNQONQ--- 378
QY 681 ILFHHILRVBPLLKIRSAQKQVQDYFYQIFMEKNEKVGVPYTIQQLL-EWSFINSNLKFA 739
Db 379 -----NQNN-----QNNQATSTPSNLSVLSKPPALNLSKIV 412
QY 740 EAPSCILIQMPRGK-DFKLFRK-----IPFSLNITDLEDTPRQCRIQGLAMVSECRE 794
Db 413 NTPPKNIKSPKTKGAVDEKSYLNAENIFFILK--VKDLRSIT-TTTTIIMVLELYHRR- 468
QY 795 CYDDPDISAGKIQFCKTCTQVHLHPKRLNHNKYNP 830
Db 469 -----QTHLLPRRLVMVNP 483

RESULT 4

US-09-262-537-58
; Sequence 58, Application US/09262537
; Patent No. 6479256
; GENERAL INFORMATION:
; APPLICANT: Havlick, Joel
; TITLE OF INVENTION: Lectomedin Materials and Methods
; FILE REFERENCE: 27866/35307
; CURRENT APPLICATION NUMBER: US/09/262,537
; CURRENT FILING DATE: 1999-03-04
; EARLIER APPLICATION NUMBER: 60/076,782

EARLIER FILING DATE: 1998-03-04
NUMBER OF SEQ ID NOS: 64
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 58
LENGTH: 1469
TYPE: PRT
ORGANISM: Homo sapiens
US-09-262-537-58

Query Match 2.5%; Score 125.5; DB 4; Length 1469;
Best Local Similarity 20.7%; Pred. No. 0.075;
Matches 130; Conservative 80; Mismatches 220; Indels 199; Gaps 35;
QY 172 GGGFTDGVYQKQFCDGDCGVALDKLELIEDDDTALESYAGGDTWQVELPPEIN 231
DB 208 GTGFTV--VYDGAVALYKERTRNIVKYD-----WGKTDIDLAVDENGWLW 232
QY 232 SRVSLKGGTIESGTIVFCVDLPKESGLGVFGVDMNDPIGNWGRFD-----GVLC 283
DB 233 LTRIKSGETV-INTANYHDTSPYR-----WGKTDIDLAVDENGWLW 274
QY 284 SPACVESTILLHINDIIPESVTOE-----RRPKLAFMSRG-----VGDKGSS 326
DB 275 IYATEGNGRLVVSQNLNPTLRFEGTWETGYDKRSASNAFMVCGVLYLRSVYVDDSEA 334
QY 327 SHNKPKATGSTDPGNRRSELFTYTLNG-----SSVDSQPSKSKNTW--YIDEVAEDPAK 379
DB 335 AGNRVDYAFTN--ANREPVSLTFPNPYQFISVVDYNDPRDNLQYVNNYF-----383
QY 380 SLTEISTDP--DRSSPPLQPPVNSLTENRHSLSFSLTKMPNTNGSIHSPLSLSAQ 436
DB 384 -VVYSLFEGFPDPSAGPATSPPLSTTTA---RPTPLTSTASPAATPLRRAPLTHPV 439
QY 437 SVMEELN-----TAPVQSPPLAMPNGSHGLEVGSIAEVENKPPFYGVIRWIGOPGL 490
DB 440 GAINQLGPLDPATAP--PSTRPPAPNLHVSPELFCPEPR-----VRRVQWP---486
QY 491 NEVLAGELEDEACAGCTDGTGRTYFTCALKKALF---VKLKSCRDPDSRFASLQPVSN 546
DB 487 -ATQOQMLVE---RPCPKGT-RGIASFQCLPALGLWNPGRPDLSNC-----TSPWVN 533
QY 547 QI-----ERCNSLA-----FGYLSVVEENTPPKMEKEGLEIMIGKKGIQ 588
DB 534 OVAQKIKSGENANIASELARHTRGSIYAGDVSSV-----KLMQLDILDAQLALRP 587
QY 589 -----GHYNSCYLDSLTLFCIFAFSSVLDTV--LLRPKEKNDVEYVSETQELLRTI 638
DB 588 PIERESAGKNYKMKHREBT-CKDYIKAVVETVDNLLRP-----EALSWKDMNATEQV 640
QY 639 NPLRIYGVYCATKIMKRLKILEKVEAASGFTSEKDPPEFLNLFHHLRVEPLLKIRSA 698
DB 641 H-----TATMLL---DVLE--EGAFLLADNVREPARFLAAKENVVLEVTVL---NTE 684
QY 699 GQVQDCVYQO-IFMEKNE-KVGVPITQ 725
DB 685 GQ-VQELVFPQBEYPRKNSIQLSAKTIQ 712

RESULT 5

US-09-262-537-20
Sequence 20, Application US/09262537
Patent No. 6479256
GENERAL INFORMATION:
APPLICANT: Haylick, Joel
TITLE OF INVENTION: Lectocidin Materials and Methods
FILE REFERENCE: 27866/35307
CURRENT APPLICATION NUMBER: US/09/262, 537
CURRENT FILING DATE: 1999-03-04
EARLIER APPLICATION NUMBER: 60/076, 782
EARLIER FILING DATE: 1998-03-04
NUMBER OF SEQ ID NOS: 64
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 20

LENGTH: 1466
TYPE: PRT
ORGANISM: Rattus rattus
US-09-262-537-20

Query Match 2.4%; Score 118.5; DB 4; Length 1466;
Best Local Similarity 20.6%; Pred. No. 0.35;
Matches 117; Conservative 80; Mismatches 207; Indels 163; Gaps 31;
QY 233 RVSLKGGTIESGTIVFCVDLPKESGLGVFGVDMNDPIGNWGRFD-----GVLC 284
DB 234 RTRIKSGETV-INTANYHDTSPYR-----WGKTDIDLAVDENGWLW 275
QY 285 PACVESTILLHINDIIPESVTOE-----RRPKLAFMSRG-----VGDKGSS 327
DB 276 YATEGNGRLVVSQNLNPTLRFEGTWETGYDKRSASNAFMVCGVLYLRSVYVDDSEA 335
QY 328 HNKPKATGSTDPGNRRSELFTYTLNG-----SSVDSQPSKSKNTW--YIDEVAEDPAK 380
DB 336 GNRVDYAFTN--ANREPVSLAFPNPYQFVSSVDYNDPRDNLQYVNNYF-----383
QY 381 LTRISTDF--DRSSPPLQPPVNSLTENRHSLSFSLTKMPNTNGSIHSPLSLSAQ 437
DB 384 VVYSLFEGFPDPSAGPATSPPLSTTTA---RPTPLTSTASPAATPLRRAPLTHPV 440
QY 438 VMEELN-----TAPVQSPPLAMPNGSHGLEVGSIAEVENKPPFYGVIRWIGOPGLN 491
DB 441 AINQLGPLDPATAP--APSTRPPAPNLHVSPELFCPEPR-----VRRVQWP---486
QY 492 EVLAGELEDEACAGCTDGTGRTYFTCALKKALF---VKLKSCRDPDSRFASLQPVSN 547
DB 487 ATQOQMLVE---RPCPKGT-RGIASFQCLPALGLWNPGRPDLSNC-----TSPWVN 534
QY 548 I-----ERCNSLA-----FGYLSVVEENTPPKMEKEGLEIMIGKKGIQ- 588
DB 535 VAQKIKSGENANIASELARHTRGSIYAGDVSSV-----KLMQLDILDAQLALRP 588
QY 589 -----GHYNSCYLDSLTLFCIFAFSSVLDTV--LLRPKEKNDVEYVSETQELLRTI 639
DB 589 IERESAGKNYKMKHREBT-CKDYIKAVVETVDNLLRP-----EALSWKDMNATEQV 641
QY 640 PLRIYGVYCATKIMKRLKILEKVEAASGFTSEKDPPEFLNLFHHLRVEPLLKIRSA 699
DB 642 -----TATMLL---DVLE--EGAFLLADNVREPARFLAAKENVVLEVTVL---STE 684
QY 700 QKQDCVYQOIF-MEKNEKVGVPITQ 725
DB 685 GQVQELVFPQBEYASESIQLSANTIK 711

RESULT 6

US-08-811-519-1
Sequence 1, Application US/0811519B
Patent No. 6630345
GENERAL INFORMATION:
APPLICANT: Petrenko, Alexandre
TITLE OF INVENTION: CALCIUM INDEPENDENT RECEPTOR OF ALPHA-LATROTOXIN,
TITLE OF INVENTION: CHARACTERIZATION AND USES THEREOF
FILE REFERENCE: 1049-1-007
CURRENT APPLICATION NUMBER: US/08/811,519B
CURRENT FILING DATE: 1997-03-04
NUMBER OF SEQ ID NOS: 31
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 1471
TYPE: PRT
ORGANISM: rat
US-08-811-519-1

Query Match 2.4%; Score 118.5; DB 4; Length 1471;
Best Local Similarity 20.6%; Pred. No. 0.35;
Matches 117; Conservative 80; Mismatches 207; Indels 163; Gaps 31;

[illegible]

Query Match	2.3%;	Score 118;	DB 3;	Length 267;
Best Local Similarity	34.3%;	Pred. No. 0.019;		
Matches	46;	Conservative 16;	Mismatches 54;	Indels 18; Gaps 8;
Qy	217	PGDTMQVELPPLINRSVSLKGGTETTESGVIFCDVLPKGESLGYFVGVDMDNPIGNWDG	276	
Db	104	PGNIMLSAL-GLRLGRVLDGQKT---GTLRFCGTT--EFASGQWGVGVELDPEGKNKG	157	
Qy	277	RFDGV---LCS-----FACVESTILLHINDIIPESVYQERRPPKLFAMSRGVGDKGSSSH	328	
Db	158	SVGGVRVYFICPKQGLFASVSK--VSKAVDAPSSVSTSTPTPRMDF-SRVTG-KGRRH	213	
Qy	329	NKPKATGSTSDPGN	342	
Db	214	KGKKKSPSPSLGS	227	

```

RESULT 8
US-09-350-614-57
; Sequence 57, Application US/09350614
; Patent No. 6689581
; GENERAL INFORMATION:
; APPLICANT: Rhodes, Kenneth
; APPLICANT: Betty, Maria
; APPLICANT: Ling, Huai-Ping
; APPLICANT: An, Wenqian
; TITLE OF INVENTION: POTASSIUM CHANNEL INTERACTORS AND USES THEREFOR
; FILE REFERENCE: WNI-070CP
; CURRENT APPLICATION NUMBER: US/09/350,614
; CURRENT FILING DATE: 1999-07-09
; EARLIER APPLICATION NUMBER: USSN 09/298,731
; EARLIER FILING DATE: 1999-04-23
; EARLIER APPLICATION NUMBER: USSN 60/110,277
; EARLIER FILING DATE: 1998-11-30
; EARLIER APPLICATION NUMBER: USSN 60/110,033
; EARLIER FILING DATE: 1998-11-25
; EARLIER APPLICATION NUMBER: USSN 60/109,333
; EARLIER FILING DATE: 1998-11-20
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 57
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Rattus sp.
US-09-350-614-57

```

Query Match	2.3%	Score 118;	DB 4;	Length 267;
Best Local Similarity	34.3%;	Pred. No. 0.019;		
Matches 46;	Conservative 16;	Mismatches 54;	Indels 18;	Gaps 8;
Qy	217	PGDTMQVLEPLLEINSRVSLKGGETIESGTVIFCDVLPGKESLGYFVGVDMDNFIQNWGD	276	
Db	104	PGNLMLSAL-GLRLGDRVLIDGQKT-----GTURFCGTT--EFASQWVGVELDEPEGKNDG	157	
Qy	277	RFDGV---LCS-----FACVESTILLHINDIIPESVTQERRPPPKLAFMSRGVGDKGSSSH	328	
Db	158	SVGGVRVFI CPPKQGLFASVSK--VSKAVDAPSSSVTSTPTPRMDF-SRVYTG-KGRREH	213	
Qy	329	NKPKATGSTSDPGN	342	
Db	214	KGKKKSPSSPSLGS	227	

RESULT 9
US-09-270-767-45699
; Sequence 45699, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270.767

; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 45699
; LENGTH: 677
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-45699

Query Match 2.3%; Score 118; DB 4; Length 677;
Best Local Similarity 20.5%; Pred. No. 0.1;
Matches 81; Conservative 59; Mismatches 154; Indels 102; Gaps 18;

QY 196 ALDKLELIEDDOTALSDVAGPGDTMVELPPPLEINSRVSLKGGTIESGTIVFCDVLPG 255
||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
DB 226 ALESITMCDDGTLIKDSHIGIDEISLTISK-----TAGGCSTMONST----- 269
QY 256 KESLGYFVGVDMDNPIGNWDGRFDGVLCSPAC--VESTILL-----HINDIIPES 303
::: : ::: : ::: : ::: : ::: : ::: : ::: : ::: : ::: :
DB 270 -DSMSITPLAADRTMP-----VVVASCPALQRTMILGEEMIGDTTFNLVDSLTTLS 318
QY 304 VTQERRPPKLAFMRSGVGDGKSSSHNKPATGTSIDPGNRRSELFPYTLINGSSVDSPQSK 363
||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
DB 319 ALQSE-----SESLPVDGNATFKRPTASATTADETQ-----VLTG-----R 354
QY 364 SKNTWVIDEVAEDPAKSILTISTDFDRSSPQLQPPPVPVNSLTENRFHSLPFSLTKMPNTN 423
||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
DB 355 QNWTFTDGC--NTPGRCETPENIDRKALLTWESTPLTNNRSHCHYHN-----NNN 407
QY 424 GSIGHSPLSLSAQSVMEELNTPVQESPPPLAMPFGNSHGLVGSIAEVKENPPFYGVIRW 483
||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
DB 408 NKAGVTPTLKRG---DMNLSPVIGATP-QKPTGTAPGRLLNNTPEPVAKTAPFNGE--- 459
QY 484 ICQPGNLNEVLGLEDEDECAGCTDGT-----RGTRYFTCALKKALFVKLKS CRPDSRF 538
||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
DB 460 -----KFVLDTMELLEQTQEPLDGTYNLQMSEQHRMQCMVDLA-----BAEVEM 504
QY 539 ASLQPVSNOIERCNSLAFFGGYLSEVVEENTPPPKMEK 574
||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
DB 505 LAQQGDEEQFE--NMLAELGKVNTLNEEQL--RMQK 536

RESULT 10
US-09-248-796A-17565
; Sequence 17565, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; TITLE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 17565
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (21)
; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unknown
US-09-248-796A-17565

Query Match 2.3%; Score 117; DB 4; Length 386;
Best Local Similarity 18.8%; Pred. No. 0.046;
Matches 75; Conservative 63; Mismatches 162; Indels 100; Gaps 17;

```

Qy 194 FVALDKLELEHDDDDTALESDDYAGPGDTMQVELPPL-EINSRVSLKGGETIESGTVIFCDV 250
Db 2 FTTFPEIQIMSSSSSSLSLSSXTATTATTSARIRLPSISELTSRSTISGGSNNGSALKSQI 61
Qy 253 LPKESLGIFYGVGDMNDPIGNWGRFGVLCSFACVESTILLHINDIIIPESVTQERRPPK 312
Db 62 SPRLSDDTSRILPSILKNTSGS-----ST-----FTSSSTPFKCPP 96
Qy 313 LAFMSRGVGDKGSSSHNKPRAT-GTSDPGNRRSELFYTLNGSSVDSOPQSKSKNTWID 371
Db 97 I-----KSTVGCTLSSGNTQSN--YVLGNTKINSILPRLSSPTLPDAKV 136
Qy 372 EVAED-----PAKSLTEISTDFDRSSPPLQPPPVNSLITEN-----RFHSLPF----- 414
Db 137 PQOQPHLPASSLSPVTRVI--NTPPQPOSVSASTSPNTQYQYVYQQQSSPIQQQQQ 194
Qy 415 -----SLTKPNTNGSTGH--SPLSLSAQSVMEELNTAPQBSRPLAMPNGNSHGLEV 465
Db 195 QOQATPAATPTVMQAOQPSHPAPLOYATQOQYPO-----FVYQSPAGVPPPPSPVTHQ 250
Qy 466 GSIAEVKENPFYGVIRWIGOPPCGLNEVLAGLEL-----EDSCAGCTDCTF-----R 512
Db 251 GHIAVHQHGHLPQGVNGMPPNVGTYIQQPEIWNKSTNRCHRC--GTTETPEWRRGPK 308
Qy 513 GTRYF--TCALKKALFVKLSCRPDSRFASLQPVYSNOIER 550
Db 309 GVRTLCNACGLFHAQVK---RKGAAALAAEEVLNNKVK 344

RESULT 11
US-08-222-617A-8
; Sequence 8, Application US/08222617A
; Patent No. 5882879
; GENERAL INFORMATION:
; APPLICANT: Veenstra, Annemarie E.
; APPLICANT: Martin, Juan F.
; APPLICANT: Garcia, Bruno D.
; APPLICANT: Gutierrez, Santiago
; APPLICANT: Barredo, Jose L.
; APPLICANT: Von Doehrens, Hans
; APPLICANT: Palissa, Harriet
; APPLICANT: Van Liempt, Henk
; APPLICANT: Montenegro, Eduardo P.
; TITLE OF INVENTION: A Method for Influencing Beta-Lactam
; TITLE OF INVENTION: Antibiotic Production and for Isolation of Large
; TITLE OF INVENTION: Quantities of ACV Synthetase
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/222,617A
; FILING DATE: 04-APR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; REFERENCE/DOCKET NUMBER: 97,157
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 798 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein

```

; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Bacillus brevis
US-08-222-617A-8

Query Match 2.3%; Score 117; DB 2; Length 798;

Best Local Similarity 18.5%; Pred. No. 0.17;
Matches 145; Conservative 109; Mismatches 273; Indels 256; Gaps 35;

QY 57 HSRIPSAKQKQIGLKILEQHAFLVDEDDVV-----EINEKTELLLAITNCE 106
| : : : : :
DB 18 HALVYPAQK-----SIHQLEFEQAEAPPRVAIVFENRRLSYQELNRKANQALRAL----- 69
| : : : : :
QY 107 ERFSLFKNRNLKGLQID--VGCPCVKQLRSGEEKFPFGRPLLAERTVSGIFFGV 164
| : : : : :
DB 70 -----LEKGVQTDTSIVGVMM-----EKSIENVI-----AILAVLRAGGAYVPI 107
| : : : : :
QY 165 E-----LLEBGRGGFTDGVVQKQLFCQDEDCGF-----VALDKLEIEDDDTAL 210
| : : : : :
DB 108 DIEYPRDRIQYILQDSQ----TKVLTQKSVQSLVHDVGYSGEVVVLDE-EQLDARETAN 162
| : : : : :
QY 211 ESDYAGPGDTMOVELPPLPINSRVSLKGGETIESGTVFICD-----VLPGKESLGYFV 263
| : : : : :
DB 163 LHQPSKPTLAV-----IYTGTTGPKMTLMLEHKGIAICNPFSKIRLASPKSTGSGFLP 218
| : : : : :
QY 264 GYDMONPIGNWGRFDGVLCSFACVBEITILLHINDIIPESVTQERRPPPKLAFMSRGVGDK 323
| : : : : :
DB 219 ACSTHFPCK-----CSWLCCLA-----PRVHPSKQIHF 249
| : : : : :
QY 324 GSSSHNKPKATGSDPDGNRRSELFTYLNQSSVDSQPSQSKNTWYIDEVAEDPAKSLTE 383
| : : : : :
DB 250 AAFEH-----YLSNELATITLPT-----YLTLTLPERTISLRI 284
| : : : : :
QY 384 ISTDFRSPPLPPPPVNSLTENRPHSLPFSLTMPNTNGSIGHSPLSLSAQSVMEELN 443
| : : : : :
DB 285 MITAGSASAPL-----VNKKDKLRIN-----AYGTETISICATIWEAPSN 327
| : : : : :
QY 444 TAPVQSPPLAMPPPGNHSHGLEVSLAEVKENPPFYGVIRWIGOPPGLNEVLGLGLELEDEC 503
| : : : : :
DB 328 QLSVQ-SVPIGKPIQNT-----IYIV-----NEDLQLLPTADEG 361
| : : : : :
QY 504 AGCTDGTGRTGYFTCALKKALFVKLSCRPDSRPFASLPVSNQIERNCSLAFGGYLSEV 563
| : : : : :
DB 362 ELCIGVGLARGYN-----RPD--LTAERFVDNPFVPGKMYRTGLAKW 405
| : : : : :
QY 564 VRENTPPKMEKGLEIM--IGKKKGIOGHYNSCYLSTLFCFLAFSSVLDTVLLRPEKN 621
| : : : : :
DB 406 LTDGT-----IEFLGRIDHQVKIRGH--RIELGEIESVLLAHEHTEAVVIAREDOH 455
| : : : : :
QY 622 DVE-----YYSETQELLRTEIVNPLRIY-----GYVCATKIMKLRKI----- 658
| : : : : :
DB 456 AGYLCAYISQAEATPAQ-----LRDYAAQLPAYMLPSYFVKLDKRNPLTPNDKIDRKAL 511
| : : : : :
QY 659 -----LEKVAASGFTSEKDPBFLNLFPHILRVEPLKIR-----SAGQKVQ 703
| : : : : :
DB 512 PEPDLTANQSAAYHPRTETESILVSMQNVLGIEK-IGIRDNFYSLGSDSIQAQVVA 570
| : : : : :
QY 704 DCYFYQIFMEKNEKGVPTIQQLLEWSFNLSNPKFAE-----APSLIIQMPRF 753
| : : : : :
DB 571 RLHSYQLKLETKDLLNYPTIEQVA--LFVKSTTRKSDQGIAGNVPLTP-----IQKWFPG 624
| : : : : :
QY 754 KDF 756
| : : : : :
DB 625 KNF 627
| : : : : :

RESULT 12

US-09-949-016-7766
; Sequence 7766, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: fastseq for Windows Version 4.0
; SEQ ID NO 7766
; LENGTH: 719
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7766

Query Match 2.3%; Score 115.5; DB 4; Length 719;

Best Local Similarity 23.3%; Pred. No. 0.19;
Matches 78; Conservative 34; Mismatches 128; Indels 95; Gaps 18;

QY 292 ILLHINDIIPESVTQERRPPKLAFLMSRGVCDKG--SSSH-----NPKATGSDPDGNRRS 345
| : : : : :
DB 26 LLLH-----PLEATHLQSP-----RQKGEADVSSPHFCEPNVPKGLADRKQNDQRK- 73
| : : : : :
QY 346 ELFYTLNGSSVDSQPSQSKNTWYIDEVAEDPAKSLTEISTDFRSPPLPPPPVNSLT 405
| : : : : :
DB 74 ----VSQGLRAPPP-----PVEKSEIAIEQKENFDLPQY----- 106
| : : : : :
QY 406 ENRFLSPFLSLTKMPNTNGSIG-HSPLSLSAQSVM-----BELNTAPVQES--- 450
| : : : : :
DB 107 ----ETTPKGLAPVTNSSGKMAIINSPQGPVSELSGKQLKTGWGSPPLPRESPTQDAAGV 162
| : : : : :
QY 451 -PPLAM---PPGNHGLEVGSIAEVK---ENPPFYGVIRWIGQP-----PGLNEVLGLLEL 499
| : : : : :
DB 163 GPPASQGRGAPGAPMGPEAGSKAELPPTVSRPPLRLGLSWDSGPEEPGRLOKVLAKLPL 222
| : : : : :
QY 500 EDS---CAGTDCGTFR---GTRYFTCALKKALFVKLSCRPDSRF-----ASLQPV 545
| : : : : :
DB 223 ABEEKFPAGKAGKAKAPGLKDFQVQVPRMOKLTKLREHILMRNQNVLGKLPDLS 282
| : : : : :
QY 546 NQIERCNSLAFGGYLVSEVVEENTPPKMEKEGLEIM 580
| : : : : :
DB 283 EAAEQEKGLP--SELSPAIEE-----EESKSGLDVM 311
| : : : : :

RESULT 13

US-09-248-796A-17555
; Sequence 17555, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 17555
; LENGTH: 523
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-17555

Query Match 2.3%; Score 114; DB 4; Length 523;
Best Local Similarity 21.2%; Pred. No. 0.15;
Matches 90; Conservative 39; Mismatches 151; Indels 144; Gaps 17;

Db 1748 SANITVEILPDEDPEDLKAFSVSVSSGSLGAHINAILTVLASDDPYGIFIFPEKNRP 1807
QY 271 IG-----NWDGRFDGVLCSFACVESTILLHINDIIPESVTQERR---PPKLA 314
Db 1808 VKVEEATQNTLSITELKGLMGKVLVSATLDA-----MEKPPYFPFNLA 1852
QY 315 FMSRG---VCDKSS--SHNKRKATGSTS---DPCNRRSELYTLNGSSVDSQPQSKS- 364
Db 1853 RATQGRDYIPASGFALFGANQSEATIAISILDDDEPERSESFVIELNLTAKVQSRSI 1912
QY 365 -----KNTWYIDEVAEDPAKSLTEISTDFDRSSPPLOPPPVNSLTENRPHSLPF 414
Db 1913 PNSPLRGPVKVETIAQIIIIANDDAFGTLOLSAPIVRAENHVGPPIINVTRTGGAFADSV 1972
QY 415 SLTKMPTNGSICHSPLSLSAQSVM-----EELNTAPVQESPPPLAMPNGNSH----- 461
Db 1973 KFKAVPIT--ATAGEDYSIASSGVVLEGETSKAVPIYVINDIYPPELGESFLGQLMNETT 2030
QY 462 -GLEVGS LAE-----VKENPPFYGV-----I 481
Db 2031 CGARLGALTEAVIIIEASDDPYGLFGFQITKLIVEPEFNSVKVNLPIIRNSGTLGNVTV 2090
QY 482 RWI-----GO-----PQC-----LNEVLG-----LELEDECAGC 506
Db 2091 QNVATINGQATGDLRWVSGNVTFAPGETIQTLLLEVLADDVPEIBEVIQVQLTDASGGG 2150
QY 507 TDGTFRGTRYFTCALKKALFVKLCKRPSRPFASLPQVSNQIERCNSLAFGGYLSEVVEE 566
Db 2151 TIG-----LDRINIIIPANDDPYGTVAFAQVY 2179
QY 567 NTPPKMEKEGLEIMIGKKIGIOGHYNSCYLDSLFLCLFAPSSVLDTVLLRPKEKNDVEYY 626
Db 2180 RVQPELRSYANITVRRSG--GHFGR-----LLL FYSTSDIDVVALAMEEQDLLSY 2230
QY 627 SET-----QELLRT-----EIVNPLRIYGVCA TKMKILEKVEAASGTSEEKDP 675
Db 2231 YESPIQGVDPDPLWRWTWNVSAVGEPL-----YTCATILCK-----EQACSAFSPFSASEGP 2281
QY 676 EEF 678
Db 2282 QRF 2284

Search completed: April 18, 2005, 15:26:33
Job time : 59 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 18, 2005, 15:33:50 ; Search time 376 Seconds
(without alignments)
4129.858 Million cell updates/sec

Title: US-09-671-687A-3
Perfect score: 5034
Sequence: 1 MSSGLWSQKVTSPYWEERI.....RLLCDAYMCYQSPWLSLYK 949

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV=xlp

-MODEL=frame-p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09671687/runat_18042005_115019_17155/app_query.fasta_1.1095
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DLOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09671687 @CGN 1 1 105 @runat_18042005_115019_17155 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:*
1: /cgn2_6/prodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/prodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/prodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/prodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/prodata/1/ina/PCUTUS_COMB.seq.*
6: /cgn2_6/prodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3676	73.0	2523	US-09-620-312D-290	Sequence 290, Appl
2	2468	49.0	2116	US-09-646-403-3	Sequence 3, Appl
3	185	3.7	110	US-09-513-999C-20085	Sequence 20085, A
4	151	3.0	5857	US-09-220-132-79	Sequence 79, Appl
5	139	2.8	2825	US-09-949-016-1895	Sequence 1895, Ap
6	139	2.8	3130	US-09-949-016-5779	Sequence 5779, Ap
7	135	2.7	2643	US-09-399-913-56	Sequence 56, Appl
8	135	2.7	2643	US-09-350-614-56	Sequence 56, Appl
9	133	2.6	3245	US-09-774-528-352	Sequence 352, App
10	130.5	2.6	1452	US-09-248-796A-10072	Sequence 10072, A
11	128.5	2.6	5610	US-09-262-537-57	Sequence 57, Appl
12	126	2.5	1503	US-08-999-774A-11	Sequence 11, Appl

13	126	2.5	4282	4	US-09-976-594-799	Sequence 799, Appl
14	126	2.5	6474	4	US-09-949-016-1717	Sequence 1717, Ap
15	122.5	2.4	1463	4	US-09-270-767-15085	Sequence 15085, A
c 16	122.5	2.4	31826	4	US-09-902-540-1256	Sequence 1256, Ap
17	121.5	2.4	13977	3	US-09-484-970B-60	Sequence 60, Appl
c 18	121.5	2.4	24333	3	US-09-639-207-9	Sequence 9, Appl
19	119.5	2.4	3545	3	US-08-885-291-54	Sequence 54, Appl
20	119.5	2.4	3545	3	US-09-496-672-54	Sequence 54, Appl
21	119.5	2.4	5715	3	US-09-107-847-1	Sequence 1, Appl
22	119	2.4	580073	4	US-08-545-528D-1	Sequence 1, Appl
23	118.5	2.4	5391	4	US-08-811-519-2	Sequence 2, Appl
24	118.5	2.4	5693	4	US-09-262-537-19	Sequence 19, Appl
25	118	2.3	1380	3	US-09-339-303-1	Sequence 1, Appl
26	118	2.3	2033	4	US-09-270-767-14130	Sequence 14130, A
27	118	2.3	2418	4	US-09-614-221A-495	Sequence 495, App
c 28	118	2.3	16047	4	US-09-902-540-1136	Sequence 1136, A
29	118	2.3	36820	4	US-09-949-016-16665	Sequence 16665, A
30	117.5	2.3	1794	4	US-09-949-016-2105	Sequence 2105, Ap
31	117.5	2.3	4926	2	US-08-853-310-1	Sequence 1, Appl
32	117	2.3	1158	4	US-09-248-796A-3462	Sequence 3462, Ap
33	117	2.3	2729	1	US-08-412-431-2	Sequence 2, Appl
34	117	2.3	2729	1	US-08-623-679-2	Sequence 2, Appl
35	117	2.3	2729	3	US-08-933-774-2	Sequence 2, Appl
36	117	2.3	2729	3	US-09-181-030-2	Sequence 2, Appl
37	117	2.3	2729	3	US-09-534-242-2	Sequence 2, Appl
38	117	2.3	2729	3	US-09-454-854-2	Sequence 2, Appl
39	117	2.3	2729	3	US-09-164-671-2	Sequence 2, Appl
40	117	2.3	2729	4	US-09-182-113-2	Sequence 2, Appl
41	117	2.3	2729	4	US-08-862-442-2	Sequence 2, Appl
c 42	116.5	2.3	1859	4	US-09-489-039A-3406	Sequence 3406, Ap
43	115.5	2.3	6671	1	US-08-280-443-1	Sequence 1, Appl
44	115.5	2.3	6671	1	US-08-457-459-1	Sequence 1, Appl
45	115.5	2.3	6671	1	US-08-555-678-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-09-620-312D-290
; Sequence 290, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunging
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_Fl_genes version 1.0
; SEQ ID NO 290
; LENGTH: 2523
; TYPE: DNA

3 later

ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (81)..(2138)
US-09-620-312D-290

Alignment Scores:

Pred. No.: 0 Length: 2523
Score: 3676.00 Matches: 692
Percent Similarity: 99.28% Conservative: 0
Best Local Similarity: 99.28% Mismatches: 3
Query Match: 73.02% Indels: 2
DB: 4 Gaps: 2

US-09-671-687a-3 (1-949) x US-09-620-312D-290 (1-2523)

QY 255 GlyLysGluSerLeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrp 274
DB 45 GGAAATAAAGCTTTTATATTTTGTGTGGACATGGATAACCCATTTGGCAACTGG 104
QY 275 AspGlyArgPheAspGlyVal--LeuCysSerPheAlaCysValGluSerThrIleLeu 293
DB 105 GATGAAGATTGATGGAGTGCAGCTTTGTAGTTTTCGTGTGGTGAAGTACAAATCTA 164
QY 294 LeuHisIleAsnAspIleIleProGluSerValThrGlnGluArgProProLysLeu 313
DB 165 TTGCACATCAATGATATCATCCAGAGAGTGTGACGAGAAAGGAGGCTCCCAACTT 224
QY 314 AlaPheMetSerArgGlyValGlyAspLysGlySerSerSerHisAsnLysProLysAla 333
DB 225 GCCTTTATGTCGAAGAGTGTGGGACCAAGGTTTCATCCAGTCATAATAAACAAGGCT 284
QY 334 ThrGlySerThrSerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsn 352
DB 285 ACAGGATCTACCTCAGACCCCTGGAATAAGAACAGATCTGAATATTTTATACCTTAAT 344
QY 353 GlySerSerValAspSerGlnProGlnSerLysSerLysAsnThrPrpTyrIleAspGlu 372
DB 345 GGGTCTCTGTGACTCACACACACCAATCCAAATCAAAAATACATGTTGATGATGAA 404
QY 373 ValAlaGluAspProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSer 392
DB 405 GTTGCAAGAGACCTCGAAATCTCTTACAGAGATATCTACAGACTTTGACCGTCTTCA 464
QY 393 ProProLeuGlnProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeu 412
DB 465 CCACCACTCCAGCTCTCTCTGAACTCCTGAACTCCTGACCCAGAACAGATTCCTCTTA 524
QY 413 ProPheSerLeuThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSer 432
DB 525 CCATTCAGTCTCACCAAGATGCCATACCAATGGAAATTTGGCCACAGTCCACTTTCT 584
QY 433 LeuSerAlaGlnSerValMetGluGluLeuAsnThrAlaProValGlnGluSerProPro 452
DB 585 CTGTACGCCAGTCTGTAATGGAAGAGCTAAACACATGCACCCGCTCCAAAGAGTCCACCC 644
QY 453 LeuAlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLys 472
DB 645 TTGGCCATGCCCTCTGGGAACCTCACATGGTGTAGAAAGTGGGCTCATTTGGCTGAAGTTAAG 704
QY 473 GluAsnProProPheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGlu 492
DB 705 GAGAACCTCTCTTATGGGTAAATCCGTGTGATCGGTGAGCCAGCACGAGACTGAATGAA 764
QY 493 ValLeuAlaGlyLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArg 512
DB 765 GTGCTCGCTGGACTGGAACCTGGAAGATGAGTGTGCAGGCTGTACGGATGGAACCTTCAGA 824
QY 513 GlyThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPheValLysLysSerCys 532
DB 825 GGCACCTCGGTATTTACCTGTGCCCTGGAAGAGGCGCTGTTTGTGAACACTGAAGAGCTGC 884
QY 533 ArgProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsn 552

DB 885 AGGCCTGACTCTAGGTTTCATCATTCAGCCGGTTTCCAAATCAGATTTGAGCCCTGTAAC 944
QY 553 SerLeuAlaPheGlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMet 572
DB 945 TCTTTAGCAATTTGGAGGCTACTTAAGTGAAGTAGTAGAAGAAATACTCCACCAAAATG 1004
QY 573 GluLysGluGlyLeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsn 592
DB 1005 GAAAAAGAGGCTTGCAGATAATGATTGGGAAGAAAGGCAATCCAGGGTCAATCAAT 1064
QY 593 SerCysTyrIleuAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThr 612
DB 1065 TCTTGTACTTACTAGACTCAACCTTATTTCTGCTTATTTTCTGCTTTTCTGTCGACACT 1124
QY 613 ValLeuLeuArgProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeu 632
DB 1125 GTGTACTTACTAGACCCAAAGAAAGACGATGTAGAATATTTATAGTGAACCCAGAGCTA 1184
QY 633 LeuArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIle 652
DB 1185 CTGAGGACAGAAATTTCTTAATCTCTGAGAATATATGGATATGTGTGTGCCACAAAAATT 1244
QY 653 MetLysLeuArgLysIleLeuGluLysValGluAlaAspSerGlyPheThrSerGluGlu 672
DB 1245 ATCAAACTCAGGAAAAATCTTGAAGAGGTGGAGGCTGCATCAGGATTTACCTCTGAAGAA 1304
QY 673 LysAspProGluGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeu 692
DB 1305 AAAGATCTCTGAGGAATTTCTTGAATATTTCTGTTTCATCATATTTTAAAGGGTAGAACCTTTG 1364
QY 693 LeuLysIleArgSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMet 712
DB 1365 CTAAAAATTAAGATCAGCAGGTCAAAAGGTACAAAGATTGTACTTCTATCAAAATTTTATG 1424
QY 713 GluLysAsnGluLysValGlyValProThrIleGlnGlnLeuLeuGluTrpSerPheIle 732
DB 1425 GAAAAAATCAGAAAGTTGGCGTTCCCAAAATTCAGCAGTTGTTAGAATGCTCTTTTATC 1484
QY 733 AsnSerAsnLeuLysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPhe 752
DB 1485 AACAGTAACCTGAAATTTGCAGAGGACCATCATGCTGATTTTCAGATGCCTCGATTT 1544
QY 753 GlyLysAspPheLysLeuPheLysIlePheProSerLeuGluLeuAsnIleThrAsp 772
DB 1545 GGAAGAGACTTTAAACTATTTAAAAAATTTTCTCTCTGGAATTTAAATATAACAGAT 1604
QY 773 LeuLeuGluAspThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCys 792
DB 1605 TTACTTTGAAGACACTCCAGACAGTCCGGATATGTGGAGGCTTGCATATGATGAGTGT 1664
QY 793 ArgGluCysTyrAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThr 812
DB 1665 AGAAGATGCTTACACCATCCGACATCTCAGCTGGAAAAATCAAGCAGTTTGTAAACCC 1724
QY 813 CysAsnThrGlnValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSer 832
DB 1725 TGCAACACTCAAGTCCACCTTCATCCGAAGAGGCTGAATCATAAATATAACCCAGTGTCA 1784
QY 833 LeuProLysAspLeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMet 852
DB 1785 CTTCCCAAGACTTACCCGACTGGGACTGGAGACACGGCTGCATCCCTTGCAGAAATG 1844
QY 853 GluLeuPheAlaValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGly 872
DB 1845 GAGTTATTTGCTGTTCTCTGATAGAAACAAAGCCACTATGTTGCTTTTGTGAAGTATGGG 1904
QY 873 LysAspAspSerAlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsn 892
DB 1905 AAGGACGATTTCTGCTGGCTCTTCTTTTGACAGCATGSCCGATCGGGATGTTGGTGTGAG 1964
QY 893 GlyPheAsnIleProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSer 912

Db 1965 GGCTTCAACATTCCTCAAGTCACCCCATGCCAGAGTAGGAGACTTGAAGATGTCT 2024
Qy 913 LeuGluAspLeuHisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeu 932
Db 2025 CTGGAGAGACCTGCATTCCTTGGACTCCAGAGAAATCCAGAGGCTGTGCACGAAGACTGCTT 2084
Qy 933 CysAspAlaTyrMetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
Db 2085 TGTGATGATATATGTGCTATGTACAGAGTCCAAACAAATGAGTTGTACAAA 2135

RESULT 2
US-09-646-403-3
; Sequence 3, Application US/09646403
; Patent No.: 6734174
; GENERAL INFORMATION:
; APPLICANT: WALLACH, David
; APPLICANT: KOVALENKO, Andrei
; TITLE OF INVENTION: MODULATORS OF THE FUNCTION OF RECEPTORS OF THE TNF/NGF RECEPTOR
; TITLE OF INVENTION: AND OTHER PROTEINS
; FILE REFERENCE: WALLACH=27
; CURRENT APPLICATION NUMBER: US/09/646,403
; CURRENT FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: IL 123758
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: PCT/IL99/00158
; PRIOR FILING DATE: 1999-03-18
; PRIOR APPLICATION NUMBER: IL 126024
; PRIOR FILING DATE: 1998-09-01
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 2116
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (691)..(691)
; OTHER INFORMATION: n is unknown.
US-09-646-403-3

Alignment Scores:
Pred. No.: 1,28e-266 Length: 2116
Score: 2468.00 Matches: 524
Percent Similarity: 91.21% Conservative: 5
Best Local Similarity: 90.34% Mismatches: 33
Query Match: 49.03% Indels: 22
DB: 4 Gaps: 6

US-09-671-687A-3 (1-949) x US-09-646-403-3 (1-2116)

Qy 387 AspPheAspArgSerProProLeuGlnProProProValAsnSerLeuThrThrGlu 406
Db 15 GACTTTGACCGTCTTTCACACCACTCCAGCCCTCCTCTGTGACTCATGTACCACCGAG 74
Qy 407 AsnArgPheHisSerLeuProPheSerLeuThrLysMetProAsnThrAsnGlySerIle 426
Db 75 AACAGATTCCTCTTACCATTCAGTCTCACCAGATGCCCAATACCAATGGAAGTATT 134
Qy 427 GlyHisSerProLeuSerLeuSerAlaGlnSerValMetGluLeuAsnThrAlaPro 446
Db 135 GGCACAGTCCACTTCTCTGTGAGCCCTCTGTAATGAAGAGTAAACACTGCACCC 194
Qy 447 ValGlnGluSerProProLeuAlaMetProProGlyAsnSerHisGlyLeuGluValGly 466
Db 195 GTCCAAGAGAGTCCACCTTGGCCATGCTCTCTGGGAATTCACATGGTCTAGAAGTGGGC 254
Qy 467 SerLeuAlaGluValLysGluAsnProProPheTyrGlyValIleArgTirpIleGlyGln 486
Db 255 TCATTGGCTGAAGTTAAGAGAACCTCTCTTCTATGGGGTAATCCGTTGGATCGGTGAG 314
Qy 487 ProProGlyLeuAsnGluValLeuAlaGlyLeuGluLeuGluAspGluCysAlaGlyCys 506
Db 315 CCACCAGGACTGAATGAAGTGTGCTGCTGAGTGGAACTGGAAGTGAAGTGTGCTGAGGCTGT 374

Qy 507 ThrAspGlyThrPheArgGlyThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPhe 526
Db 375 ACGGATGGAAACCTTTCAGAGCACTCGGTATTTTCACTGTGCCCTGAAAGAGCGCTGTTT 434
Qy 527 ValLysLeuLysSerCysArgProAspSerArgPheAlaSerLeuGlnProValSerAsn 546
Db 435 GTGAAACTGAAGAGCTGCAGGCTGACTCTAGGTTTGCATCATTTGCAGCCGGTTTCCAAT 494
Qy 547 Gln-11eGluArgCysAsnSerLeuAlaPheGlyGlyTyrLeuSerGluValValGluG1 566
Db 495 CAAAGATTGAGCGCTGTAACTCTTTTAGCATTTGGAGGCTACTTAAAGTGAAGTAGT-GAAGA 553
Qy 566 uAsnThrProLysMetGluLysGluGlyLeuGluIleMetIleGly----LysLysLys 585
Db 554 AAATACTCCACCAAAAATGGAAAAGAAAGCTTGGAGATAATGATGGGGAAGAAAGAAA 613
Qy 585 sGlyIleGlnGlyHisTyrAsnSerCysTyrLeuAspSerThrLeuPhe-CysLeuPhe- 604
Db 614 GGCATCCAAGGGTCATTACAAATCTTGKTACTTAGACTCAACCTTATCTKGCTTATTTK 673
Qy 605 AlaPheSerSerValLeu-AspThr---ValLeuLeuArgProLysGluLysAsnAsp-- 622
Db 674 GCTTTTAGTTCGTCTCTGAGCACTGTGTCTTACTTTTAGACCCCAAGAAAAGAAACGAT 733
Qy 623 -ValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGluIleValLeuProLeuAr 642
Db 734 GTTAGAATATTTWTWKWMMMACCAAGAGCTACTGAGGACAGAAATGTGTAATCTCTGAG 793
Qy 642 gIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArgLysIleLeuGluLysVa 662
Db 794 AATATATGATATGTGTGTGCCCAAAAATTTAGAAACTGAGGAAAATACTTGAAGAGGT 853
Qy 662 lGluAlaAlaSerGlyPheThrSerGluGluLysAspProGluGluPheLeuAsnIleLe 682
Db 854 GGAGGCTGCATCAGGATTTACCTCTGNAGAAAAGATCCTGAGGAAATCTTGAATATCT 913
Qy 682 uPheHisHisIleLeuArgValGluProLeuLeuLysIleArgSerAlaGlyGlnLysVa 702
Db 914 GTTTCATCATATTTTAAAGGTAGAACCTTTGCTAAAAATAAGATCAGCAGGTCAAGAGGT 973
Qy 702 lGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGluLysValGlyValProTh 722
Db 974 ACAAGATTTGTACTTCTATCAAAATTTTATGAAAAAAATGAGAAAGTTGGCGTTCCAC 1033
Qy 722 rIleGlnGlnLeuLeuGluTrpSerPheIleAsnSerAsnLeuLysPheAlaGluAlaPr 742
Db 1034 AATTGACAGTGTGTTAGATGGTCTTTTATCAACAGTAACTGAAATTTGCAGAGGCACC 1093
Qy 742 oSerCysLeuIleIleGlnMetProArgPheGlyLysAspPheLysLeuPheLysIle 762
Db 1094 ATCATGCTGATTTATTCAGATGCTCGATTTGGAAAAGACTTTAAACTATTTAAAAAATT 1153
Qy 762 ePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAspThrProArgGlnCysAr 782
Db 1154 -TTTCCTTCTCTGGAATTAGATATAACAGATTTTACTTGAAGACAC-CCCAGACAGTGC 1211
Qy 782 gIleCysGlyLysLeuAlaMetTyrGluCysArgGluCysTyrAspAspProAspIleSe 802
Db 1212 GATATGAGAGGCTGTGCAATGATGAGTGTA- GAATGCTACGACCATCCGACACCCAG 1270
Qy 802 rAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGlnValHisLeuHisProLy 822
Db 1271 C---TGAAAAACAAGCAGTTTTGTAAAACTGCAACACTCAAGTCCACCTTCATCCGA 1327
Qy 822 sArgLeuAsnHisLysTyrAsnProValSerLeuProLysAspLeuProAsp---TrpAs 841
Db 1328 GAGGCTGAATCATATAATATAAACCCAGTGTCACTTCCCAAGACTTACCCGACTGGAGA 1387
Qy 841 pTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAlaValLeuCysIleG1 861
Db 1388 TTGGAGACGCGGTGTCATCCTTTGCCAGAAATATGGAGTTATTTGCTGTTCTCTCATAGA 1447

Qy 861 uThSerHisTyrValAlaPheValLysTyrGlyLysAspSerAlaThrLeuPhePhe 881
Db 1448 AACAGCCACTATGTTGTTTGAAGTATGGAAGACGATTCCTGCTGCTCTTCTT 1507
Qy 881 e-AspSerMetAlaAsp-ArgAspGlyGlyGlnAsnGlyPheAsnIleProGlnValThr 900
Db 1508 TGGACAGCATGCCATCCGGATGGTGTGAGATGCTCAACATTCCTCCCAAGTCMCC 1567
Qy 901 ProCysProGluValGlyGlnTyrLeu-LysMetSer-LeuGluAsp---LeuHisSerL 919
Db 1568 CMTGCCAGAGTAGGAGTAGTCTTGGAGATGCTCTCTGGAAGACCTGSAWTCCTT 1627
Qy 919 euAspSerArgArgIle-GlnGlyCysAlaArgArgLeuLeuCysAspAla-TyrMetCy 938
Db 1628 GGACTCCAGGAGAAATCCAGGCTGTGCACGAGACTGCTTTGTGATGCCATATATGTG 1687
Qy 938 s-MetTyr-GlnSerProThrMetSerLeuTyrLys 949
Db 1688 CCATGTACCCAGAGTCCCAACATGATGATTTGTACAAA 1723

RESULT 3

US-09-513-999C-20085
; Sequence 20085, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 20085
; LENGTH: 110
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-513-999C-20085

Alignment Scores:
Pred. No.: 4,2e-12 Length: 110
Score: 185.00 Matches: 36
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.68% Indels: 0
DB: 4 Gaps: 0

US-09-671-687A-3 (1-949) x US-09-513-999C-20085 (1-110)

Qy 595 TyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeu 614
Db 1 TACTTAGACTCAACCTTATTCGTTATTGCTTTTGTCTTTAGTCTGTTCTGGACACTGTGTTA 60
Qy 615 LeuArgProLysGlnLysAsnAspValGluTyrTyrSerGluThrGln 630
Db 61 CTTAGACCCCAAGAAAAGAACGATGATGATATATTATAGTGAACCCAA 108

RESULT 4

US-09-220-132-79
; Sequence 79, Application US/09220132
; Patent No. 6506607
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT
; TITLE OF INVENTION: OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 07334-074001
; CURRENT APPLICATION NUMBER: US/09/220,132
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: US 60/079,303

; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: US 60/068,821
; PRIOR FILING DATE: 1997-12-24
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 79
; LENGTH: 5857
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-220-132-79

Alignment Scores:

Pred. No.: 4,31e-05 Length: 5857
Score: 151.00 Matches: 107
Percent Similarity: 31.58% Conservative: 61
Best Local Similarity: 20.11% Mismatches: 184
Query Match: 3.00% Indels: 180
DB: 4 Gaps: 23

US-09-671-687A-3 (1-949) x US-09-220-132-79 (1-5857)

Qy 219 AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly 238
Db 277 GAGACTCAGGAGGAATTTGTGATGACTTTCGAGTTGGGAGGCGAGTTTGGTGAATGGA 336
Qy 239 -----GlyGluThrIleGluSerGlyThrValIlePhe 249
Db 337 AATAAGCCTGGATTATCCAGTTTCTTGAGAAACC----- 372
Qy 250 CysAspValLeuProGlyLysGluSerLeuGlyTyrPheValGlyValAspMetAspAsn 269
Db 373 ---CAGTTTGCACAGCCAG-----TGGCTGGAAATTTTGTAGATGAA 414
Qy 270 ProIleGlyAsnTrpAspGlyArgPheAspGlyValLeuCysSerPheAlaCysValGlu 289
Db 415 CCATAGGCAAGAACGATGCTCGGTGGCAGGAGTTCGG---TATTTCCAGTGTGAACCT 471
Qy 290 SerThrIleLeuLeuHisIleAsnAspIleIleProGluSerValThrGlnGluArg 309
Db 472 -----TTAAAGGCGCATATT-----ACCGA 492
Qy 310 ProProLysLeuAlaPheMetSerArgGlyValGlyAspLysGlySerSerHisAsn 329
Db 493 CCTTCAAAGTTA-----ACAAGGAGGTG-----CAAGCA 522
Qy 330 LysProLysAlaThrGlySerThrSerAspProGlyAsnArgArgSerGluLeuPheTyr 349
Db 523 GAAGATGAAGCTAATGGCTTGCAGACACACCGCCCTCCCGAGCTACTTCACCGCTGTGC 582
Qy 350 ThrLeuAsnGlySerSerValAspSerGlnProGlnSerLysSerLysAsnThrTrpTyr 369
Db 583 ACTTCTACGGCCAGCATGGTGTCTTCTCCCTCCACCCCTTCAAC----- 630
Qy 370 IleAspGluValAlaGluAspProAlaLysSerLeuThrGluIleSerThrAspPheAsp 389
Db 631 ATCCCTCAGAAACCATCACAGCCAGCA----- 657
Qy 390 ArgSerSerProProLeuGlnProProProValAsnSerLeuThrThrGluAsnArgPhe 409
Db 658 ---GCAAGGAACCTTCAGCTACGCTCCGATCAGCAACCTTACA----- 699
Qy 410 HisSerLeuProPheSerLeuThrLysMetProAsnThrAsnGlySerIleGlyHisSer 429
Db 699 ----- 699
Qy 430 ProLeuSerLeuSerAlaGlnSerValMetGluLeuAsnThrAlaProValGlnGlu 449
Db 700 -----AAAACTGCCAGTGAATCTATCTCAACCTTTTCAGAGGTGGCTCAATCAAGAAA 753
Qy 450 SerProLeuAlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAla 469
Db 754 -----GGAGAAAGAGAGCTCAAAATCGGAGACAGAGTA 786

```

QY 470 GluValLysGluAsnProPheTyrGlyValIleArgTrpIleGlyGlnProProGly 489
Db 787 TTGTTGGTGGCACTAAGCT-----GGTGTAGTCCGGTTCTTGGGGAGACCGACTTT 840
QY 490 LeuAsnGluValLeuAlaGlyLeuGluLeuGluAspGluCysAlaGlyCysThrAspGly 509
Db 841 GCCAAGGGGGAGTGTGTGGCGTGGAGTTA---GATGAGCCACTTGGGAAGAATGATGGC 897
QY 510 ThrPheArgGlyThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPheValLysLeu 529
Db 898 GCTGTGTGGCAACAAGGATTTTCACTGTGTCAACCCAAATATATGCTTGTTCGCT----- 951
QY 530 LysSerCysArgProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIleGlu 549
Db 952 -----CTGTCCAC 960
QY 550 ArgCysAsnSerLeuAlaPheGlyGlyTyrLeuSerGluValValGluGluAsnThrPro 569
Db 961 AAAGTTTACCAAGATTGGCTTC-----CCTTCCACTACACCA 996
QY 570 ProLysMetGluLysGluGlyLeu---GluIleMetIleGlyLysLysLysGlyIleGln 588
Db 997 GCCAAGCCCAAGCCCAACGAGTGGCGAGTGTATGCGCAGTGTATGCGCAGCTGAAG 1056
QY 589 GlyHisTyrAsnSerCysTyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSer 608
Db 1057 CCGAGCCCTTCGCTTCCTCCCTCAGCTCC-----ATGAGCTCA 1095
QY 609 ValLeuAspThrValLeuLeuArgProLysGluLysAsnAspValGluTyrTyrSerGlu 628
Db 1096 GTGGCTCTCTGTGAGCAGCAGGCC-----AGTCGG 1128
QY 629 ThrGlnGluLeuLeuArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCys 648
Db 1129 ACAGGACTATTGACTGAAACCTCTCCCTTACGCCAGGAAGATCTCCGCTACCACTGCC 1188
QY 649 AlaThrLysIleMetLys-----LeuArgLysIle 658
Db 1189 CTCAGGAGGCCCTGAGGAGGAAGCAGCAGCATGTCGTGCGGAGCAAGGAT 1248
QY 659 LeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGluGluPhe 678
Db 1249 CTGGAGAGGGCGAGGTGGCCCAAGGCC---ACGAGCCAGTGGGGAGATAGAGCAGGAG 1305
QY 679 LeuAsnIleLeu-----PheHisIleLeu----- 687
Db 1306 CTAGCTCTGGCCCGGAGCAGCATGACAGCATGCTCTGGAATTGGAAGCCAAATGGAC 1365
QY 688 -----ArgValGluProLeuLeuLysIle 695
Db 1366 CAGCTCGCAACAATGTGGAAGCTGCTGACAGGAGAGGTGGAGCTTCTCAACCAGCTT 1425
QY 696 ArgSerAlaGlyGlnLysValGlnAspCysTyrPhe 707
Db 1426 GAAGAGGAGAAAGGAAGTGTGAGGACCTTCAGTTC 1461

```

RESULT 5

```

US-09-949-016-1895
; Sequence 1895, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08

```



```

; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1895
; LENGTH: 2825
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-1895

Alignment Scores:
Pred. No.: 0.000251      Length: 2825
Score: 139.00           Matches: 131
Percent Similarity: 33.06%      Conservative: 74
Best Local Similarity: 21.13%   Mismatches: 232
Query Match: 2.76%             Indels: 184
DB: 4                       Gaps: 28

```

US-09-671-687A-3 (1-949) x US-09-949-016-1895 (1-2825)

```

QY 51 GlnAspArgSerValGlyHisSerArgIleProSerAlaLysGlyLysAsnGlnIle 70
Db 17 CAAGCCAGAGAGCCGGAGGAAATGAAATTCCTCAATGCC----- 55
QY 71 GlyLeuLysIleLeuGluGlnProHisAlaValLeuPheValAspGluAspValValGlu 90
Db 56 -----CTCTCTAGCTGGTTGACCTCCGCTGTGTGCAATC----- 94
QY 91 IleAsnGluLysPheThrGluLeuLeuAlaIleThrAsnCysGluGluArgPheSer 110
Db 95 -----TCAAGCCGGCTCTCA 109
QY 111 LeuPhe-----LysAsnArgAsnArgLeuSerLysGlyLeuGlnIleAspVal 126
Db 110 CTGTGGCTTGTGCTCCGGGAAATGGAACCCAGCTTGGGACAGGATGGGAATGATTTG 169
QY 127 GlyCysProValLysValGlnLeuArgSerGlyGluGluLysPheProGlyValValArg 146
Db 170 ACTGTGCTTTGGCGTCTCCCGAGCCCTGTGGAGCCAG----- 208
QY 147 PheArgGlyProLeuLeuAlaGluArgThrValSerGlyIlePhePheGlyValGluLeu 166
Db 209 -----GCCTCTTGGAGCATCTTTGGGGCTCAGCGCA 238
QY 167 LeuGluGluGlyArgGlyGlnGlyPheThrAspGlyValTyrGlnGlyLysGlnLeuPhe 186
Db 239 GCGGAGGTTCGGGCGACACACGTGGCCACTCCCGCAG----- 274
QY 187 GlnCysAspGluAspCysGlyPheValAlaLeuAspLysLeuGluLeuGluAspAsp 206
Db 275 -----GAGGCTGCC 283
QY 207 AspThrAlaLeuGluSerAspTyrAlaGlyProGlyAspThrMetGlnValGluLeuPro 226
Db 284 ATGCCCCACATTCGGAGGACGAGGAGCCCGCGAGAGCCACAGGAGCCAGGCCCT 343
QY 227 Pro-----LeuGluIleAsnSerArg-Valse 235
Db 344 GCGGCCCAAGTAAGTCTCTTACCGCAGGAGTATCTTGAGTCCAACTCCACGATTGT 403
QY 235 rLeuLysGlyGlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGly 255
Db 404 CTGACTGGGGATGCCACTTTCACCAAGAGGAGAAAC-----GACAA-AAAACCTGG 453
QY 255 yLysGluSerLeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAs 275
Db 454 CCAACAGAGT-----TCACAGTCCCAAGAGGCTTTCTC 489
QY 275 pGlyArgPheAspGlyValLeuCys-SerPheAlaCysValGluSerThr-----IleL 293
Db 490 ACCGACACTTGAAGTGTCCACTGCTCCCTCATCTGTGGACCCCGGGGGCACATCA 549
QY 293 euLeuHisIleAsnAspIleIleProGlu---SerValThrGlnGluArgProProL 312
Db 550 TTGACCTGGTGAATGATCCAGCTGCCAGACATCATGATCTCAGAGGAGGACAAAGAGAAA 609

```


Db 574 CTCAACAGCTCCGTGAAGACCGCAACAGAGTCGGATCCAACTCTCAGACAGCGGCTCT 633
QY 471 VallysGluAsnProProPheTyr----- 478
Db 634 GTGAAGCGGGCGAAAGAGACCTGCGCCTGGGGACCGGCTGCTGGTGGCGGAGCAAG 693
QY 479 ---GlyValileArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGlyLeu 497
Db 694 ACTGGCGTGGTGGTGGTGGGAGACAGACTTTGCCAAGGGCGAGTGGTGGCGGIG 753
QY 498 GluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyrPhe 517
Db 754 GAGCTG---GACGAGCCCTTGGGAGAAATGATGGGCGGTGGCGGACACAGTACTTC 810
QY 518 ThrCysAlaLeuLysLysAlaLeuPheValLysLeu----- 529
Db 811 CAGTGGCCACCAAGTTTGTCTCTTCGCGGCCATCCAAAGTGATCGGTATCGGCTTC 870
QY 530 -----LysSerCysArgProAspSerArgPheAla----- 539
Db 871 CCATCTACAGCCCGCCAGCCAGGCAAGCAAGAACCAAGCGGTATGGCCATGGGTGTGTCAGCA 930
QY 540 ---SerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAla-----Phe 556
Db 931 CTGACCCACAGTCCACAGAGTTCCTCCATCAGCTCCGTGAGTCTGTGGCTCTCCGTC 990
QY 557 GlyGlyTyrLeuSerGlu-----ValValGluGluAsnThrProProLysMetGlu 573
Db 991 GGGGTGGCCCGCAGCGCAGTGGCTGCTCACGAGACCTTTCACGCTAGCGCCGCAAG 1050
QY 574 LysGluGlyLeuLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 593
Db 1051 APTCTGGGACACCGCCCTTGCAGGAGGACCTGAAGGAGAGCAGCAGCAGCAGCAGCAG 1110
QY 593 rCysTyrLeuAspSerThr-----LeuPheCysLeuPheAlaPheSerSerValLe 610
Db 1111 CTGCTGGCTGAACGAGACTGGACCGGCTGAG-GTGGCCAGGCCACAGCCACATCTG 1169
QY 610 uAspThr-----ValLeuLeuArgProLysGluLysAsnAspValGluTyrTy 626
Db 1170 CGAGGTGGAGAGGAGATTGCCCTGTCTCAAGCAGCAGCATGAG-----CAGTATGT 1220
QY 626 rSerGluThrGlnGlu---LeuLeuArgThrGluLysValAlaAsnProLeuArgIleTyrGl 645
Db 1221 TGCAGAAAGCCGAGGAGAGCTGAGGAGCGCCGCTGCTC----- 1260
QY 645 yTyrValCysAlaThrLysLysLysLysLysLysLysLysLysLysLysLysLysLys 665
Db 1261 -----GTGAGAGCGTGGCGGAAA-----GAGAAGTGGACCTGTC 1295
QY 665 aSerGlyPheThrSerGluGluLysAspProGluGlu 677
Db 1296 CAACAGCTGGAGGAGGAGGAGGAGGAGTGGAGGAT 1332

RESULT 7

US-09-399-913-56
; Sequence 56, Application US/09399913
; Patent No. 6361971
; GENERAL INFORMATION:
; APPLICANT: Rhodes, Kenneth
; APPLICANT: Betty, Maria
; APPLICANT: Ling, Huai-Ping
; APPLICANT: An, Wendian
; TITLE OF INVENTION: POTASSIUM CHANNEL INTERACTORS AND USES THEREFOR
; FILE REFERENCE: MNI-070CP2
; CURRENT APPLICATION NUMBER: US/09/399,913
; CURRENT FILING DATE: 1999-09-21
; EARLIER APPLICATION NUMBER: USN 60/110,277
; EARLIER FILING DATE: 1998-11-30
; EARLIER APPLICATION NUMBER: USN 60/110,033
; EARLIER FILING DATE: 1998-11-25
; EARLIER APPLICATION NUMBER: USN 60/109,333
; EARLIER FILING DATE: 1998-11-20

; EARLIER APPLICATION NUMBER: USN 09/298,731
; EARLIER FILING DATE: 1999-04-23
; EARLIER APPLICATION NUMBER: USN 09/350,614
; EARLIER FILING DATE: 1999-07-09
; EARLIER APPLICATION NUMBER: USN 09/350,874
; EARLIER FILING DATE: 1999-07-09
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 56
; LENGTH: 2643
; TYPE: DNA
; ORGANISM: Rattus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(801)
US-09-399-913-56
Alignment Scores:
Pred. No.: 0.00627 Length: 2643
Score: 135.00 Matches: 131
Percent Similarity: 31.65% Conservative: 51
Best Local Similarity: 22.78% Mismatches: 172
Query Match: 2.68% Indels: 222
DB: 3 Gaps: 30
US-09-671-687A-3 (1-949) x US-09-399-913-56 (1-2643)
QY 169 GluGlyArg-----GlyGlnGlyPheThrAspGlyValTyrGlnGlyLysGln 184
Db 201 CAAGGACAGGACCGCTGGTGGCCAAAGAAATTCGCGACGCTGTAGNAGAGGC----- 254
QY 185 LeuPheGlnCysAspGlu-AspCysGlyPheValAlaLeuAspLysLeuGluLeuIleGl 204
Db 255 -----TGTCACACTGTCCTGC-----ACCTTCCTAAAGTCACACTACCCAA 296
QY 204 uAspAspAspThrAlaLeuGluSerAspTyrAlaGlyProGlyAspThrMetGlnValGl 224
Db 297 CTATGCAACAGTC-----CCAGGCAATCTCATGCTCAGCGC 332
QY 224 uLeuProProLeuGluIleAsnSerArgValSerLeuLysGlyGlyGluThrIleGluSe 244
Db 333 GCTG---GGCTCGCTCTAGGACCGAGTGTCTCTCGATGGCCAGAGAGC----- 381
QY 244 rGlyThrValIlePheCysAspValLeuProGlyLysGluSerLeuGlyTyrPheValGl 264
Db 382 -GGCAGCTGAGGTCTCTGGGGACCACT-----GAGTTCGACAGTGGCCAGTGGTGG 434
QY 264 yValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPheAspGlyVal----- 281
Db 435 CGTGGAGCTAGATGAACCGGAGGCAAGACGACGCGGAGCGTGTGGGGGTGTCCGCTACTT 494
QY 282 -LeuCysSer-----PheAlaCysValGluSerThrIleLeuLeuHisI 296
Db 495 CATCTGCCCTCCCAAGCAGGGTCTTTTGTATCTGTGTCCAG-----GTCCTCAAGGC 548
QY 296 eAsnAspIleIleProGluSerValThrGlnGluArgProProLysLeuAlaPheMe 316
Db 549 AGTGGATGCACCCCTCATCTGTACTCCACGCCCGCAGCTCCCGGATGGACTTC-- 606
QY 316 tSerArgGlyValGlyAspLysGlySerSerSerHisAsnLysProLysAlaThrGlySe 336
Db 607 -TCCCGTGTAAACGGGC---AAAGCGCGAGGGAACACAAA----- 642
QY 336 rThrSerAspProGlyAsnArgSerGluLeuPheTyrThrLeuAsnGlySerSerVa 356
Db 643 -----GGGAAGAAAGATCCCA----- 660
QY 356 lAspSerGlnProGlnSerLysSerLysAsnThrTyrIleAspGluValAlaGluAs 376
Db 660 ----- 660
QY 376 pProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerProProLeuGl 396
|||||


```
Db 607 -TCCCGTGTAACGGGC---AAAGCGCGGAGGAACACAAA----- 642
Qy 336 rThrSerAppProGlyAsnArgArgSerGluLeuPheTyrThrLeuAsnGlySerSerVa 356
Db 643 -----GGGAAGAAGAGTCCCA----- 660
Qy 356 lAspSerGlnProGlnSerLysSerLysAsnThrTyrTyrIleAspGluValAlaGluAs 376
Db 660 ----- 660
Qy 376 pProLalysSerLeuThrGluIleSerThrAspPheAspArgSerSerProProLeuG1 396
Db 661 -----TCTTCCCA----- 669
Qy 396 nProProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLe 416
Db 669 ----- 669
Qy 416 uThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaG1 436
Db 670 -----TCTCTGGGC-----AGCCTGCA 686
Qy 436 n-SerValMetGluGluLeuAsnThrAlaProValGlnGlnSerProProLeuAlaMetP 456
Db 687 GCAGCGTG-----AAGGGGCCAAAGCTGAAG--- 712
Qy 456 roProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProp 476
Db 713 -----TTGGAGACCAAGTCTGTGGCAGGCGCAGAACAA----- 745
Qy 476 roPheTyrGlyValIleArgTyrIleGlyGlnProProGlnGlnSerLeuAsnGluValLeuAlaG 496
Db 746 -----GGGATTGGCGTTCTATGGGAAGACAGACTTTGCTCCAGGTTACTGGTATG 797
Qy 496 lLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyPheArgGlyThrArgT 516
Db 798 GCATTGAAGT---GACCAGCCACCGGCAAGCATACGGCTCTGTGTCGGTCCGGT 854
Qy 516 yPheThrCysAlaLeuLysLysAlaLeuPheValLysLysSerCysArgProAspS 536
Db 855 ACTTTTACCTGTGCCCGGAGCAGCGGGTCTTTGCA-----CCAGCAT 896
Qy 536 erArgPhe-----AlaSerLeuGlnProValSerAsnGlnIleGluArgCysA 552
Db 897 CTCGTATCCAGAGGATTGGTGGATCCACTGATCCCTCGAGACAGTGTGGAGCAAA-A 955
Qy 552 snSerLeuAlaPheGlyGlyTyrLeuSerGluValValGluAsnThrPro---ProL 571
Db 956 AAAGTGCAATCAAGTGAATGACATGACACAGCCCAACGACCTTCAACAGTCCGGACCCCA 1015
Qy 571 yMetGluLysGluGlyLeuGluIleMetIleGlyLysLysGlyIleGlnGlyHis- 590
Db 1016 AA-----GGACATT 1024
Qy 591 --TyrAsnSerCysTyrLeuAspSerThrLeuPheCysLeuPheAlaPhe----- 606
Db 1025 GCATCAGAACTCTATCTCCAGGTACTCTTCTGCTGCTGTTCTCTGGATGCTGAGG 1084
Qy 607 -----SerSerValLeuAspThrValLeu----- 614
Db 1085 GCGGAGATCGATGCTTTAGAGACCTGGATACCTGACACAGACAGAGTCCCTCTAGCAT 1144
Qy 615 -----LeuArgPro-----LysG 619
Db 1145 CTCCTTGACACAGGAGACCCCGTACCTTAAGATAGAGATCCCGAGTGACACTCCAGA 1204
Qy 619 lLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGluIleValA 639
Db 1205 ATAGAACCCTGTTAGCCAGCCCTCGATTACTAGGTCCCATTTATTAACAGATCTCCCAT 1264
Qy 639 snProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArgLysIleL 659
Db 1265 GACGACTCCCCCAATACAGACCTCATGTTACCCCAAGAGAGATTCCCTGAGTAGCACCT 1324
```

```
Qy 659 euGluLysValGluAlaAlaSerGlyPheThrSerGluLysAspProGluGluPheL 679
Db 1325 TCAGGTAGTCCCTGTCCCC-----TACCCTCAGACGAGATTTCCTCCCAATAAACATT 1378
Qy 679 euAsnIleLeuPheHisHisIleLeuArgValGluProLeu 692
Db 1379 TTCCA-----CATCACCAAGGGATGCTGACCTCTC 1410
```

RESULT 9

```
US-09-774-528-352
; Sequence 352, Application US/09774528
; Patent No. 6743619
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Zhao, Qing A.
; APPLICANT: Yang, Yonghong
; APPLICANT: Xue, Aidong J.
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wang, Dunrui
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6743619el Nucleic Acids and
; FILE OF INVENTION: Polypeptides
; FILE REFERENCE: 802
; CURRENT APPLICATION NUMBER: US/09/774, 528
; CURRENT FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 441
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 352
; LENGTH: 3245
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (402)..(2537)
US-09-774-528-352
```

```
Alignment Scores:
Pred. No.: 0.00154 Length: 3245
Score: 133.00 Matches: 168
Percent Similarity: 33.02% Conservative: 110
Best Local Similarity: 19.95% Mismatches: 296
Query Match: 2.64% Indels: 271
DB: Gaps: 39
```

US-09-671-687A-3 (1-949) x US-09-774-528-352 (1-3245)

```
Qy 246 ThrValIlePheCysAspValLeuProGlyLysGluSerLeuGlyTyrPheValGlyVal 265
Db 119 TCGTAGCTCTGTGTCCAGTCCCTGGCAT-GATCTGATCGCTTGAAGCAGGTACC 177
Qy 266 AspMetAspAsnProIleGlyAsn-----Trp----- 274
Db 178 CGCTCGGTCTCCCCATCAGCTCCAGGGGCTCTTCAGTTGAGTGGTGAGCCGACAGAA 237
Qy 275 -----AspGlyArgPheAspGlyValLeuCysSerPheAlaCys----- 287
Db 238 GCCCGGGGTGAGCTCAAGGGCGGATGGGAGGGTCTCTGGGGGCGCTTAACCTGCGCACCGC 297
Qy 288 ---ValGluSerThrIleLeuLeuHisIleAsnAspIleIleProGluSerValThrGln 306
Db 298 GGCCATCAGCTGGCAGATCAITTCACAG-AAGTCTCTGTCTCTCTTTCATCTACAGTGC 356
Qy 307 GluArgArgProProLysLeuAlaPheMetSerArgGlyValGlyAspLysGlySerSer 326
Db 357 AGCTGTTTCCAGACGCTTGTCTCTCCTCAGGTGC-----GGGAGTGAT 398
```


327 SerHisAsnLysProLysAlaThrGlySerThrSerAspProGlyAsnArgSerGlu 346
Db : : : : :
339 CTGATGGCCAGGTTTCGAAACTTCTTGGCCCTCGGCTCGGTCGCGC----- 449
Qy : : : : :
347 LeuPheTyrThrLeuAsnGlySerSerValAspSerGlnProGlnSerLysSerLysAsn 366
Db : : : : :
450 -----TGAATCTCCGAGGTGGGAGGAGGAGCTCTCTCTGAGGAGCGGTGAGAG 500
Qy : : : : :
367 ThrTyrPThrIleAspGluValAlaGluAspProAlaLysSerLeuThrGluIleSerThr 386
Db : : : : :
501 GCGGGGAAATGAGAGGCGCGGCGGCGGCTACGAAGCGCTCTTCGAGAGCTGAAGCC 560
Qy : : : : :
387 AspPheAspArgSerSerProProLeuGln-----ProProValAsnSerLeu 403
Db : : : : :
561 GAG---GAGATCAAGCTGGAGCCATTACAGAGCGTGAGCCCGCGCGGAGGAGAACTTG 617
Qy : : : : :
404 Thr-----ThrGluAsnArgPheHisSerLeu-ProPheSerLeuTh 417
Db : : : : :
618 ACGTGGAGCAGCGCGGCGGAGCAGGAAGGTGCTCCCTTCAATCCCGCTTCGCTGTAC 677
Qy : : : : :
417 rlyMetProAsn-----ThrAsnGlySerIleGlyHisSerProLeuSe 432
Db : : : : :
678 AGCAGCTCTCGCCCGTTTGGCCGCGCCCAAGCCCGCTCGGCGCCGAGCCCGGGCC 737
Qy : : : : :
432 rLeuSerAlaGlnSerValMetGluGluLeuAsnThrAlaProValGlnGluSerProPr 452
Db : : : : :
738 CGCTCCGCGAGCAGCGCTGGGCTCTCGGCGCCACACC-CCGCTCCAGCGCGCCCGCC 796
Qy : : : : :
452 oLeuAlaMetProPro-----GlyAsnSerHi 461
Db : : : : :
797 CCGCGCGCCACCCCGCCCGCCACCGCGCGCCAGCGCGCTGGCGTGGATCCCGCG 856
Qy : : : : :
461 sGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProProPheTyrGlyValI 481
Db : : : : :
857 CAGATCCCGCGCGCTGGCTCCAGCGCTCAGACCGGAGAAGC-----TGCTCTGGTGACCT 910
Qy : : : : :
481 eArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGlyLeuGluLeuGluAs 501
Db : : : : :
911 AGACGGTCCGGGGATCTCGGCGCTTACGGGACTGGTCTGGAGGTGAGTTGGTCA 970
Qy : : : : :
501 pGluCysAlaGlyCysThr----- 507
Db : : : : :
971 GGGTCCACAGCGTCTCTCATGTGGAGAGCTTTAAAGTAGGTAAGAACTGGCAGAAGAA 1030
Qy : : : : :
508 -----AspGlyThr-----PheArgG 513
Db : : : : :
1031 CCTGAGGTTGATCTACCAGCGTTTCGTTGGAGTGGGACCCCGAGAGACTAGGAACGTAA 1090
Qy : : : : :
513 yThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCys-A 533
Db : : : : :
1091 AGCAAGTCATGATCTGTACATGTAGTACCCATATGAAACAGACTCCACTCTGTCT 1150
Qy : : : : :
533 rgProAspSerArgPhe---AlaSerLeuGlnProValSerAsnGlnIleGluArgCysA 552
Db : : : : :
1151 CTCCTGTGCTTTTGGCTGTCTTCACTGAGAAACATATTTCACAAACATGCAGAAACAA 1210
Qy : : : : :
552 snSer-----LeuAlaPheGlyTyrLeuSerGluValValGluG 566
Db : : : : :
1211 GCAGCACCACTTATGCTGTAGACCTTTTATCATGGGGTCATATATTGCTTGTGTGAAGA 1270
Qy : : : : :
566 lu-----AsnThrProProLysMetGluLys-GluGlyLeuGlu 578
Db : : : : :
1271 TTATGTATATGACAAAGACATAGACAGATTGCCAAGAAACAAAGAAAAATTTTGG 1330
Qy : : : : :
579 ileMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeuAspSer 598
Db : : : : :
1331 ATT-----ATTAACTTCCACTCACAGATGTTTC---TCATCA 1366
Qy : : : : :
599 ThrLeuPheCysLeu-----PheAlaPhe 606
Db : : : : :
1367 ACAGTTTATGACATCAGGTTTGAAGACAAAGCAATCAACCTGTGAGACAAAGGAACAGGA 1426

607 SerSerValLeuAspThrValLeuLeuArgProLysGluLysAsnAspValGluTyrTyr 626
Db : : : : :
1427 GCCAAATTTGGTGAAACC-----CAAGAAAAAGAGAAAAAGTCAGTCTATAC 1477
Qy : : : : :
627 -----SerGluThrGlnGluLeuLeuArgThrGluIleValAsnPro 640
Db : : : : :
1478 TGTAGGCTCGAGAGGCTAATCAATCTTGGAAACACTTGTATTATGAATTTGTATGTC 1537
Qy : : : : :
641 LeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeu-----ArgLys 657
Db : : : : :
1538 ---GGCACTTACCATTCTCTCTACTGAAGATTTCTTCTCTCTGACACGACAAATG 1594
Qy : : : : :
658 ileLeuGluLysValGluAlaAlaSerGlyPhe----- 668
Db : : : : :
1595 TATAATGACAAAGCCCGAGCTTGTCTGTGCTGTGAAATGTCTTCGCTTTTTCATGCTAT 1654
Qy : : : : :
669 -----ThrSerGluGluLysAspPro 675
Db : : : : :
1655 GTACTCTGGAGCGCAACTCTCACATTCCTATAAGTTACTGCATCTGATATGATGCCA 1714
Qy : : : : :
676 -GluGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysI 695
Db : : : : :
1715 TGCAGAACATT----- 1727
Qy : : : : :
695 eArgSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAs 715
Db : : : : :
1728 ---GCAGGCTACAGGCGAGGATGCCATGAGTCTCTTATTGCAATA----- 1772
Qy : : : : :
715 nGluLysValGlyValProThrIleGlnGlnLeuLeuGluTrpSerPheIleAsnSerAs 735
Db : : : : :
1773 ---TTAGACGTGCTACATAGACACAGCAAGATGATAGTGG 1810
Qy : : : : :
735 nLeuLysPheAlaGluAlaPro-----SerCysLeuIleIleGlnMetProArgph 752
Db : : : : :
1811 TGGCAGGAGGCCAATAAACCCCACTGCTTAAGTCTGATCATAGACCAATCTTTACAGG 1870
Qy : : : : :
752 eGlyLysAspPheLysLeuPheLysLysIlePheProSerLeuLeuLeuAsnIleThr-- 771
Db : : : : :
1871 TGGCTGCAATCAGATGTCACATGTCAGGCTGCCATAGTGTCTTCTACCAACCATAGACC 1930
Qy : : : : :
772 ---AspLeuLeuLysAspThrProArgGlnCys----- 781
Db : : : : :
1931 ATGCTGGGACATCAGTTTGACTTGGCTGCTCTGTGCCACATTCGATTCACCAAGACC 1990
Qy : : : : :
782 -----ArgIleCysGlyGlyLeuAlaMetTy 790
Db : : : : :
1991 AGAGAGGCTGACAGCACAGTACAGGAGTACCATACAGGAATCCCTCCTTAC 2050
Qy : : : : :
790 rGluCysArgGluCysTyrAspAspProAspIle-----SerAlaGlyLysIleLysG 808
Db : : : : :
2051 AGACTGTCTACAGTGGTTTACAGGCGCAGACCTAGGAGCAGTGCCTCAAAATCAAA-- 2108
Qy : : : : :
808 nPheCysLysThrCysAsnThrGlnValHis----- 818
Db : : : : :
2109 ---TGCAATAGTTGCCAAAGCTACCAAGAGTCTACTAAACAGCTCACAATGAAAAAAT 2164
Qy : : : : :
819 -----LeuHisProLysArgLeuAsnHis-----Ly 827
Db : : : : :
2165 ACCATTGTGGCTGTGTTTTCATCTCAAGCGGTTTGAGCATGTAGGCACAGAGCGGAA 2224
Qy : : : : :
827 sTyrAsnPro---ValSerLeuProLysAspLeu----- 837
Db : : : : :
2225 GATTATACCTTTATCTCTCTTCCCTTGGAGCTGGACATGACTCGGTTTTTGGCCTCTAC 2284
Qy : : : : :
838 -----ProAspTrpAspTrpArgHisGlyCysIleApr 848
Db : : : : :
2285 TAAAGAGAGCAGAATGAAGAGGCCAGCCACCAACAGAT-----TGTGTGCC 2332
Qy : : : : :
848 oCysGlnAsn---MetGluLeuPheAlaValLeu-----CysIleGluThrSe 863
Db : : : : :
2333 CAATGAGAAAGATTTCTCTTGTGAGTATTAAATCACCATGGAATCTTGGAAAGTGG 2392
Qy : : : : :
863 rHisTyrValAlaPheValLysTyrGlyLysAspSerAlaTrpLeuPhePheAsp 882

[illegible]

QY 817 ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnPro 830
 Db 1408 ACACATCTACTCCCAAGAGAAAATAGTATGATGGTGAATCCA 1449

RESULT 11

US-09-262-537-57
 ; Sequence 57, Application US/09262537
 ; Patent No. 6479256
 ; GENERAL INFORMATION:
 ; APPLICANT: Hayflick, Joel
 ; TITLE OF INVENTION: Lectomedin Materials and Methods
 ; FILE REFERENCE: 27866/35307
 ; CURRENT APPLICATION NUMBER: US/09/262,537
 ; CURRENT FILING DATE: 1999-03-04
 ; EARLIER APPLICATION NUMBER: 60/076,782
 ; EARLIER FILING DATE: 1998-03-04
 ; NUMBER OF SEQ ID NOS: 64
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 57
 ; LENGTH: 5610
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (281)...(4687)
 US-09-262-537-57

Alignment Scores:
 Pred. No.: 0.0135 Length: 5610
 Score: 128.50 Matches: 139
 Percent Similarity: 34.51% Conservative: 86
 Best Local Similarity: 21.32% Mismatches: 224
 Query Match: 2.55% Indels: 204
 DB: 4 Gaps: 37

US-09-671-687A-3 (1-949) x US-09-262-537-57 (1-5610)

QY 172 GlyGlnGlyPheThrAspGlyValTyrGlnGlyLeuPheGlnCysAspGluAsp 191
 Db 902 GGCACAGGCTTGTG-----GTCTACGATGGTGGCCGCTGTGTACAAACAGGAGCGCAG 955
 QY 192 CysGlyPheValAlaLeuAspLysLeuGluLeuIleGluAspAspThrAlaLeuGlu 211
 Db 956 CGCAACATCTCAAGTATGAC----- 976
 QY 212 SerAspTyrAlaGlyProGlyAspThrMetGlnValGluLeuProProLeuGluIleAsn 231
 Db 976 ----- 976
 QY 232 SerArgValSerLeuLysGlyGlyGluThrIleGluSerGlyThrValIlePheCysAsp 251
 Db 977 CTACGGACGCGCATCAAGACGGGGAGACGGTC---ATCAATACCGCAACTACCATGAC 1033
 QY 252 ValLeuProGlyLysGluSerLeuGlyTyrPheValGlyValAspMetAspAsnProIle 271
 Db 1034 ACCTCGCCCTACCGC----- 1048
 QY 272 GlyAsnTrpAspGlyArgPheAsp-----GlyValLeuCys 283
 Db 1049 -----TGGGGCGAAGACCGACATTGACCTGGCGGTGGACGAGACGGGCTGTGGGTC 1102
 QY 284 SerPheAlaCysValGluSerThrIleLeuLeuHisIleAsnAspIleIleProGluSer 303
 Db 1103 ATCTACGCCACTGAGGCAACAACGGCGGCTGTGTGTGAGCCAGCTGAACCCCTACACA 1162
 QY 304 ValThrGlnGlu-----ArgArgProProLysLeuAlaPhe 315
 Db 1163 CTGGCTTTGAGGCGACGTGGGAGACGGGTTACGACAGCGCTCGGCATCAACGCGCTTC 1222
 QY 316 MetSerArgGly-----ValGlyAspLysGlySerSer 326
 Db 1223 ATGTTGTGGGGTCTGTAGCTCTCGTCTCGTTGATGATGATGACAGCGGCGC 1282

QY 327 SerHisAsnLysProLysAlaThrGlySerThrSerAspProGlyAsnArgSerGlu 346
 Db 1283 GCTGGCAACCGGTGGACTATGCTTCAACACCAAT-----GCCAACCGGAGGACGCT 1336
 QY 347 LeuPheTyrThrLeuAsnGly-----SerSerValAspSerGlnProGln 361
 Db 1337 GTCAGCCTCACCTTCCCAACACCGCTACAGTTTCATCTCTCTCTGCTGACTACAAACCTTCGC 1396
 QY 362 SerLysSerLysAsnThrTrp-----TyrIleAspGluValAlaGluAspProAlaLys 379
 Db 1397 GACAACACGCTGACGCTGGAAACAATATTC----- 1429
 QY 380 SerLeuThrGluIleSerThrAspPhe-----AspArgSerSerProProLeuGln 396
 Db 1430 ---GTGGTGGCTACAGCTGGAGTTCCGGCCGCCGCCAGTGTGCGCCAGCCACT 1486
 QY 397 ProProProValAsnSerLeuThrGluAsnArgPheHisSerLeuProPheSerLeu 416
 Db 1487 TCCCAACCCCTCAGACGACGACCAACAGCC-----AGGCCACGCGCTCACCAGC 1537
 QY 417 ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGln 436
 Db 1538 ACAGCTCGCCCGCAGCCACCCCGCTCCGGCGGACCCCTCACCACGACCCAGTG 1597
 QY 437 SerValMetGluGluLeuAsn-----ThrAlaProValGlnGluSer 450
 Db 1598 GGTGCCATCAACGACGCTGGACCTGATCTGCTCCAGCCACGCCCCAGTC----- 1648
 QY 451 ProProLeuAlaMetProProGlyAsnSer-HisGlyLeuValGlySerLeuAlaGln 470
 Db 1649 CCCAGCACCGCGCGGCCCCCGACCGAATCTACAGTGTCCCTGAGCTCTCTGCG-- 1706
 QY 470 uValLysGluAsnProProPheTyrGlyValIleArgTyrIleGlyGlnProProGlyLe 490
 Db 1707 -----AGCCCGGAGAGGTACGCGG-GTCCAGTGG-----CGGCCAC 1743
 QY 490 uAsnGluValLeuAlaGlyLeuGluAspGluCysAlaGlyCysThrAspGlyTh 510
 Db 1744 CCAGCAG-----GGCATGTGTGGAG-----AGGCCCTGCCCAAGGGAC 1785
 QY 510 rPheArgGlyThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPhe----- 526
 Db 1786 T---CGAGGAATTGCTCTCTTCCAGTGTCTACACAGCTTGGGGCTCTGGAAACCCCGGG 1842
 QY 527 -ValLysLeuLysSerCysArgProAspSerArgPheAlaSerLeuGlnProValSerAs 546
 Db 1843 CCTGACCTCAGCAACTGC-----ACCTCCCTCCCTGGTCAA 1878
 QY 546 nGlnIle-----GluArgCysAsnSerLeuAla----- 555
 Db 1879 CCAGTGGCCCAAGATCAAGAGTGGGGAGAACCGCGGCAACATCCCGCAGAGTGGC 1938
 QY 556 -----PheGlyGlyTyrLeuSerGluValValGluGluAsnTh 568
 Db 1939 CCAGCACACCCGGGCTCCATCTACGGCGGGGAGCTCTCTCTCTCTGTG----- 1987
 QY 568 rProProLysMetGluLysGlyLeuGluIleMetIleGlyLysLysGlyIleGln 588
 Db 1988 -----AAGCTGATGAGCAGCTGTGTGGACATCTCTGGATGCCAGCTGCGAGCCCTGCG 2040
 QY 588 n-----GlyHisTyrAsnSerCysTyrLeuAspSerThrLe 600
 Db 2041 GCCCATCGCGGAGTCAAGCGCAAGAACTACACAAAGTGCACAGGCGAGAGAGAAC 2100
 QY 600 uPheCysLeuPheAlaPheSerSerValLeuAspThrVal-----LeuLeuArgProLys 618
 Db 2101 T---TGTAAGGATTATATCAAGCGCGGTGGAGACAGTGGACAATCTGTCTCCGCGCA-- 2155
 QY 618 sGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuArgThrGluIleVal 638
 Db 2156 -----GAAGCTCTGGAGTCTCTGGAAGGACATGAATGCCACGGAGCAGGT 2199


```
Db 898 -----AGATGGATG---AAACCCCAAAACAGAGAGATACT 933
QY 494 uLaGlyLeuGluLeuGlu 500
Db 934 CTTGGCGATGGAAGTGGAT 952

RESULT 13
US-09-976-594-799
; Sequence 799, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 799
; LENGTH: 4282
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 1385527.4
US-09-976-594-799

Alignment Scores:
Pred. No.: 0.0157 Length: 4282
Score: 126.00 Matches: 77
Percent Similarity: 34.86% Conservative: 37
Best Local Similarity: 23.55% Mismatches: 84
Query Match: 2.50% Indels: 129
DB: 4 Gaps: 21

US-09-671-687A-3 (1-949) x US-09-976-594-799 (1-4282)
QY 255 GlyLysGluSerLeuGlyTyrPheValGlyValAspMetAspAsnProLeu---GlyAsn 273
Db 126 GGCCTCGATTCTACAGTTATTAT-----GACCAGGAATTTATGTGGA 170
QY 274 TrpAspGlyArgPheAspGlyValLeuCysSerPheAlaCysValGlu----- 289
Db 171 AGTGACAGCAGATTGCTGGATACGTGACATCAATCTGCTGCAACTGAACTTGAAGATGAT 230
QY 290 -----SerThrIleLeuLeu-----His--- 295
Db 231 GACGATGACTATTCATCATCTACGAGTTCTGGTCAGAAAGCCAGGATATCATGCC 290
QY 296 -----IleAsnAspIleIleProGluSerValThrGln----- 306
Db 291 CTTGTGGCATTGCTTAATGAT---ATACCAAGTCAACAAAGATGATGATTCATTGCT 347
QY 307 GluArgArgProProLysLeuAlaPheMetSerArgGlyValGlyAspLysGlySerSer 326
Db 348 GAGCACAGACCTCCAAAGATTGCA-----GACCGG---GAAGAT 383
QY 327 SerHisAsnLysProLysAlaThr----- 334
Db 384 GAATACAAAAGCATGATGGCGGACCATGATAATTTCCCGAGCGCTTTGTATCCTTTTGGCA 443
QY 335 -----GlySerThrSerAspProGly---AsnArgArgSerGluLeu----- 347
Db 444 GATGGAGGGGAAGACCCCTGATCTTAAATGAATGCTAGGACTTACATGGATGTAATCGGA 503
QY 347 ----- 347
Db 504 GAACAACACTTGACTAAAGAAGAACGAGAAATTAGGCAACAGCTAGCAGAAAGCTAAA 563
```

```
QY 348 -----PheTyrThrLeuAsnGlySerSerValAspSerSerGlnProGlnSerLysSer 364
Db 564 GCTGGAGAACTAAAGTCGTCAATGGAGCAGCAGCG---TCCAGCGCTTCATCAAAACGA 620
QY 365 LysAsnThrTrpTyrIleAspGluValAlaGluAspProAlaLysSerLeuThrGluIle 384
Db 621 AAACGGCGTTGG-----GATCAACAGCTGATCAGACTCCTGCTGCGCACTCCCAAAAA 674
QY 385 SerThrAspPheAspArgSerProProLeuGlnProProValAsnSerLeuThr 404
Db 675 CTATCAAGTTGGGATCAGCAGAGAGACCCCTGGGCATACTCTCTTCAAGATGGGATGAG 734
QY 405 ThrGluAsnArgPheHisSerLeuProPheSerLeuThrLysMetProAsnThrAsnGly 424
Db 735 ACACCAAGGTCGT-----GCAAAGGGAAGCAGAGACTCTCTGGA 770
QY 425 SerIle-----GlyHisSerProLeuSerLeuSer 434
Db 771 GCAACCCCGAGGCTCAAAAATATGGGATCTTACACTAGCCACACACCGCGGAGCTGCT 830
QY 435 AlaGlnSerValMetGluLeuAsnThrAlaProValGlnGluSerProProLeuAla 454
Db 831 -----ACTCTGGAGCAGGTGATACACAGGCCAT 860
QY 455 MetProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsn 474
Db 861 GCGACACACAGGC-----CATGGA-----GGCGCAACTTCCAGTGTCTGTAATAAAC 905
QY 475 ProProPheTyrGlyValIleArgTyrPheGlyGlnProProGlyLeuAsnGluVal 494
Db 906 -----AGATGGGATG---AAACCCCAAAACAGAGAGATACT 941

RESULT 14
US-09-949-016-1717
; Sequence 1717, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1717
; LENGTH: 6474
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1717

Alignment Scores:
Pred. No.: 0.0337 Length: 6474
Score: 126.00 Matches: 77
Percent Similarity: 34.86% Conservative: 37
Best Local Similarity: 23.55% Mismatches: 84
Query Match: 2.50% Indels: 129
DB: 4 Gaps: 21

US-09-671-687A-3 (1-949) x US-09-949-016-1717 (1-6474)
QY 255 GlyLysGluSerLeuGlyTyrPheValGlyValAspMetAspAsnProLeu---GlyAsn 273
```

```
Db 134 GGCTCGATTCTACAGTTATTAT-----GACCAGGAAATTTATGTGGA 178
Qy 274 TrpAspGlyArgPheAspGlyValLeuCysSerPheAlaCysValGlu----- 289
Db 179 AGTGACAGAGATTTGCTGGATACGTGACATCAATGCTGCAACTGGAAGATGAT 238
Qy 290 -----SerThrIleLeuLeu-----His-- 295
Db 239 GACGATGACTATTCATCACTACAGTTGCTGGTCAGAGAAGCCAGGATATCATGCC 298
Qy 296 -----IleAsnAspIleIleProGluSerValThrGln----- 306
Db 299 CCTGTGGCATTGCTTAATGAT---ATACACAGCTCAACAGACAGTATGATCCATTGCT 355
Qy 307 GluArgArgProProLysLeuAlaPheMetSerArgGlyValGlyAspLysGlySerSer 326
Db 356 GAGCAGACAGCTCCAAAGATTGCA-----GACCGG---GAAGAT 391
Qy 327 SerHisAsnLysProLysAlaThr----- 334
Db 392 GAATACAAAGATAGCGGACCATGATAATTTCCCGAGAGCGTCTTGATCCTTTTGCA 451
Qy 335 -----GlySerThrSerAspProGly---AsnArgSerSerGluLeu----- 347
Db 452 GATGGAGGGAACCCCTGATCTTAAATGAATGCTAGGACTTACATGGATGTAATGCGA 511
Qy 347 ----- 347
Db 512 GAACACACTTGACTAAGAAAGAACAGAAATTTAGGCAACAGCTAGCAGAAAAAGCTAAA 571
Qy 348 -----PheTyThrLeuAsnGlySerSerValAspSerGlnProGlnSerLysSer 364
Db 572 GCTGGAGAACTAAAGTCGTCAATGGAGCAGCAGCG---TCCGAGCGCTCCATCAAAACGA 628
Qy 365 LysAsnThrTyrTyrIleAspGluValAlaGluAspProAlaLysSerLeuThrGluIle 384
Db 629 AAACGGCGTTGG-----GATCAACAGCTGTACAGACTCCTGGTCCCATCCCAAAAAA 682
Qy 385 SerThrAspPheAspArgSerSerProProLeuGlnProProValAsnSerLeuThr 404
Db 683 CTATCAAGTTGGATCAGGACAGAGACCCCTGGGCATACTCTTCTTAAGATGGATGAG 742
Qy 405 ThrGluAsnArgPheHisSerLeuProPheSerLeuThrLysMetProAsnThrAsnGly 424
Db 743 ACACCAAGGTCTG-----GCAAGGGAAGCGAGACTCTCTGGA 778
Qy 425 SerIle-----GlyHisSerProLeuSerLeuSer 434
Db 779 GCAACCCCGCTCAAAAATATGGGATCTTACACCTAGCCACACACAGCGGAGCTGCT 838
Qy 435 AlaGlnSerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeuAla 454
Db 839 -----ACTCTGGAGCGAGGTGATACACACAGGCCAT 868
Qy 455 MetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsn 474
Db 869 GCGACACCAAGGC-----CATGCA-----GGCGCAACTTCCAGTGTCTGTAATAAAC 913
Qy 475 ProProPheTyrGlyValIleArgTyrIleGlyGlnProProGlyLeuAsnGluValLe 494
Db 914 -----AGATGGGATG---AAACCCCAAAACAGAGAGAGATACT 949
Qy 494 uAlaGlyLeuGluLeuGlu 500
Db 950 CTGGGCGATGGAAGTGGAT 968
```

RESULT 15

```
US-09-270-767-15085
; Sequence 15085, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
```

```
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15085
; LENGTH: 1463
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-15085
```

Alignment Scores:

Pred. No.:	0.00537	Length:	1463
Score:	122.50	Matches:	69
Percent Similarity:	38.87%	Conservative:	34
Best Local Similarity:	26.04%	Mismatches:	89
Query Match:	2.43%	Indels:	73
DB:	4	Gaps:	16

US-09-671-687A-3 (1-949) x US-09-270-767-15085 (1-1463)

```
Qy 229 GluIleAsnSerArgValSerLeuLysGlyGlyuThrIleGluSerGlyThrValIle 248
Db 568 GAAATCTGTCCAGCGTG-----GGCAAGGAGAACACCAGCCCAACACGCCGTT 618
Qy 249 PheCysAspValLeuProGlyLysGluSerLeu----- 259
Db 619 TCGGCCACACACACGCGCTGGCAAAAAAATCTCATCGAACTCATCTCGGCAGGTGG 678
Qy 260 GlyTyThrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPheAsp 279
Db 679 TCGCACTTT---GGAGCAGCGTCCGACAAAGAACACCAACACCCACCCAGCAGTACGAC 735
Qy 280 -GlyValLeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsnAsp-- 298
Db 736 ATCG-----GGTGGTTCGCTGTGATACCAACGACAA 768
Qy 299 -----IleIleProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMe 316
Db 769 TCATGTAAAAAACACCAAGACCAACCAAGTATGGCCATGCTCCGCCAGAG----- 817
Qy 316 tSerArgGlyValGlyAspLysGlySerSerSerHis---AsnLysProLysAlaThrGln 335
Db 818 -TCCACGCCACATCCCCACGCGTCCAGACATATGTAATGGAATGCAATGCTCAGAATGG 876
Qy 335 ySerThrSerAspProGlyAsnArgSerGluLeuPheTyThrLeuAsnGlySerSe 355
Db 877 AGTACACGCGGA-----AATGGAAGTGG 900
Qy 355 rValAspSerGlnProGlnSerLysSerLysAsnThrTyrTrpIleAspGluValAlaGln 375
Db 901 CCAGGCGCGCAGCCCCCAGAGTGTCTAGT---AACGGATGCGACGCAAGATGCGCGAAC 957
Qy 375 uAspProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProPro-- 394
Db 958 GACTACTTTCACCTTCAACTACA-----ACCATCTCTCCAG 993
Qy 395 ----LeuGlnProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuPr 413
Db 994 GCATAGTCAAAAGTCCGCCCACTCCAATATA-----TCGCCGCC 1032
Qy 413 oPheSerLeuThrLysMetProAsnThrAsnGlySerIleGlyHisSerPro-LeuSerL 433
Db 1033 AGCTTCAGTGACA-----CCCTCTGCAATGGAATGTG---CACAGTCCGAATGCGAA 1083
Qy 433 euSerAlaGlnSerValMetGluLeuAsnThrAlaProValGlnGluSerProProL 453
Db 1084 CAGCACACCGCGCGGAAGCGGAGGAG---GAAGTAGATCTCTATACAGCAGCAACGCCGCG 1140
Qy 453 euAlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysG 473
Db 1141 TC-----AACGCTCTCAGCGTCCGAGGCTCGCGGCAACAAGGTGA 1179
```

QY 473 luAsnProPro 476
: : : : :
Db 1180 GGAGTCTCCA 1190

Search completed: April 18, 2005, 20:13:32
Job time : 425 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 15:14:54 ; Search time 73 Seconds
(without alignments)
5027.885 Million cell updates/sec

Title: US-09-671-687A-3

Perfect score: 5034

Sequence: 1 MSSGLWSQEKVTSYPWEERI.....RLLCDAYMCYQSPWLSLYK 949

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001s:*
- 5: Geneseq2002s:*
- 6: Geneseq2003as:*
- 7: Geneseq2003bs:*
- 8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4983	99.0	953	7	ADC24816
2	4983	99.0	953	8	ADQ95918 T cell ac
3	4983	99.0	960	8	ADR99244 DKFZp586D
4	4980	98.9	953	8	ADQ95916 T cell ac
5	4971.5	98.8	956	6	ABB82783 Human CYL
6	4971.5	98.8	956	8	ADRI14489 Human NF-
7	4968.5	97.7	956	8	ADQ95920 T cell ac
8	3876.5	77.0	739	4	AAU23747 Novel hum
9	3862	76.7	731	3	AAU23747 Novel hum
10	3855	76.6	731	4	AAU23747 Novel hum
11	3636	72.2	685	4	AAU23747 Novel hum
12	3563	70.8	698	4	AAU23747 Novel hum
13	2755	54.7	558	4	AAU23211 Novel hum
14	2343	46.5	476	4	AAU23211 Novel hum
15	1275	25.3	261	5	ABB89233 Human pol
16	1173	23.3	238	5	ABB89233 Human pol
17	856	17.0	517	4	ABB61669 Drosophil
18	597	11.9	113	4	AAU14965 Peptide #
19	597	11.9	113	4	AAU14965 Peptide #
20	597	11.9	113	4	AAU27399 Peptide #
21	597	11.9	113	4	AAU27399 Peptide #
22	597	11.9	113	4	AAU19377 Protein #
23	597	11.9	113	4	AAU19377 Protein #
24	597	11.9	113	4	AAU54704 Human bra
25	597	11.9	113	4	AAU02691 Peptide #

26	597	11.9	113	5	ABG36764	Abg36764 Human pep
27	568	11.3	106	8	ADK71947	Adk71947 Human ori
28	545	10.8	101	8	ADK71945	Adk71945 Human ori
29	514	10.2	96	8	ADK71949	Adk71949 Human ori
30	492	9.8	104	8	ADK71962	Adk71962 Human ori
31	491	9.8	91	8	ADK71943	Adk71943 Human ori
32	151	3.0	1392	2	AAU06999	Aay06999 Restin pr
33	151	3.0	1427	8	ADP56353	Adp56353 Human PRO
34	149	3.0	1427	7	AAU10534	Aar10534 Human 160
35	146	2.9	354	2	ABM85235	Abm85235 Human pro
36	143	2.8	547	5	AAU74342	Aau74342 Human cyt
37	143	2.8	547	5	ABU97353	Abu97353 Novel hum
38	143	2.8	547	7	ADM04066	Adm04066 Human pro
39	139.5	2.8	1921	4	ABB62962	Abb62962 Drosophil
40	136	2.7	708	8	ADN99865	Adn99865 Novel hum
41	135.5	2.7	306	3	AAU42642	Aau42642 Human ORF
42	135.5	2.7	306	7	ADB64836	Adb64836 Human pro
43	135.5	2.7	721	4	AAU87339	Aau87339 Novel cen
44	135.5	2.7	721	8	ADI54654	Adi54654 Novel hum
45	135.5	2.7	2273	6	ABU38191	Abu38191 Protein e

ALIGNMENTS

RESULT 1
ADC24816

ID ADC24816 standard; protein; 953 AA.

XX AC ADC24816;

XX DT 18-DEC-2003 (first entry)

XX DE Human breast specific polypeptide (BSP) DEX0238_137, SEQ ID NO:137.

XX KW Human; breast specific polypeptide; BSP; breast specific nucleic acid;
KW BSNAs; breast cancer; metastasis; non-cancerous disease; breast tissue;
KW identification; monitoring; diagnosis;
KW engineered breast tissue production; transgenic animal; drug screening;
KW cytostatic; gene therapy; vaccine; chromosome 16p13.3.

XX OS Homo sapiens.

XX PN WO2003020900-A2.

XX PD 13-MAR-2003.

XX PF 29-AUG-2002; 2002WO-US027777.

XX PR 31-AUG-2001; 2001US-0316306P. - *late*

XX PA (DIAD-) DIADEXUS INC.

XX PI Sun Y, Liu C, Salceda S;

XX DR WPI; 2003-290182/28.

XX DR N-PSDB; ADC24898.

XX FT New breast specific polypeptide useful for identifying, diagnosing,
PT monitoring, staging, imaging and treating breast cancer and non-cancerous
PT disease states in breast.

XX PS Disclosure; SEQ ID NO 137; 264pp; English.

XX CC The invention relates to breast specific polypeptides (BSPs) and nucleic
CC acids (breast specific nucleic acids; BSNAs) encoding them. The invention
CC also relates to vectors and host cells comprising a BSNAs sequence;
CC antibodies against BSNAs; the recombinant production of BSNAs; methods of
CC detection of BSNAs or BSNAs in a sample; kits for detecting a risk of
CC cancer or presence of cancer in a patient; and vaccines comprising a BSNAs
CC or BSP. The invention additionally discloses fragments, mutants, fusion
CC proteins, homologous proteins and allelic variants of BSNAs; methods for
CC identifying and designing agonists and antagonists of BSNAs; methods for

CC identifying and monitoring breast tissue; producing engineered breast
CC tissue for treatment and research; producing transgenic animals and cells
CC comprising BSA sequences; aptamers evolved to bind specifically to BSPs;
CC and single exon probes based on BSA sequences. BSPs, BSNAs and
CC antibodies against BSPs are useful for identifying, diagnosing,
CC monitoring, staging, imaging and treating breast cancer (including breast
CC cancer metastases), and non-cancerous disease states in breast tissue.
CC BSPs and BSNAs may additionally be used to identify and monitor breast
CC tissue, in screening for BSP agonists and antagonists, and in the
CC production of engineered breast tissue for treatment or research. BSNAs
CC may also be used in gene therapy and in the production of transgenic
CC animals and cells. The present sequence represents a breast specific
CC polypeptide (BSP) disclosed in the invention.

XX SQ Sequence 953 AA;

Query Match 99.0%; Score 4983; DB 7; Length 953;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 948; Conservative 0; Mismatches 1; Indels 4; Gaps 4;

QY 1 MSSGLWSQEKVTSYWEERIFVLLQECSTVDKQKLLKVPKSGISQYIDRSVGHSHI 60
DB 1 MSSGLWSQEKVTSYWEERIFVLLQECSTVDKQKLLKVPKSGISQYIDRSVGHSHI 60
QY 61 PSAKGGKQIGKILEQPHAVLFVDE-DVINEKFTTELLAITNCEERFSLFKNRNL 119
DB 61 PSAKGGKQIGKILEQPHAVLFVDEKDVINEKFTTELLAITNCEERFSLFKNRNL 120
QY 120 KGLQIDVGCVPKVOI RSBEKPGVVRFRGPLLAERTVSGIFGVLELLEGRGQGTGV 179
DB 121 KGLQIDVGCVPKVOI RSBEKPGVVRFRGPLLAERTVSGIFGVLELLEGRGQGTGV 180
QY 180 YGKQLFOCDEDCG-FVALDKLELTDOTALESDYAGPDQMVELPLEINRSVSLKG 238
DB 181 YGKQLFOCDEDCG-FVALDKLELTDOTALESDYAGPDQMVELPLEINRSVSLKV 240
QY 239 GETIESGTVIFCDVLPFGKESLGVFGVDMNDPIGNWDGRPDGV-LCSFACVESTILLHN 297
DB 241 GETIESGTVIFCDVLPFGKESLGVFGVDMNDPIGNWDGRPDGV-LCSFACVESTILLHN 300
QY 298 DIIIPSVTOERPPKLAFLMSRGVGDGSSSHNKPATGSTSDPGNR-RSELFYTLNGSSV 356
DB 301 DIIIPSVTOERPPKLAFLMSRGVGDGSSSHNKPATGSTSDPGNRSELFTYTLNGSSV 360
QY 357 DSQPSKSNKNTWYIDVAVSDPAKSLTEISTDFDRSSPPLOPPVNSLTENRPHSLPFSL 416
DB 361 DSQPSKSNKNTWYIDVAVSDPAKSLTEISTDFDRSSPPLOPPVNSLTENRPHSLPFSL 420
QY 417 TKMPNTNGSIGHSPLSLSAQSVMEELNTAPVQESPEPLAMPNGSHGLEVGS LAEVKENPP 476
DB 421 TKMPNTNGSIGHSPLSLSAQSVMEELNTAPVQESPEPLAMPNGSHGLEVGS LAEVKENPP 480
QY 477 FYGVIRWIGOPGLNEVLAGELEDEDCAGCTDGTFRGTRYFTCALKKALFVKLSCRPS 536
DB 481 FYGVIRWIGOPGLNEVLAGELEDEDCAGCTDGTFRGTRYFTCALKKALFVKLSCRPS 540
QY 537 RFASLOPVSNQTERCNSLAFGGVLSVRENTPPKMEKEGLEIMIGKKIGIQGHNSCYL 596
DB 541 RFASLOPVSNQTERCNSLAFGGVLSVRENTPPKMEKEGLEIMIGKKIGIQGHNSCYL 600
QY 597 DSTLFCFLFAFSSVLDVTLARPKENDEYVYSETQELLRTIEVNLPRIVGVCAKIMKLR 656
DB 601 DSTLFCFLFAFSSVLDVTLARPKENDEYVYSETQELLRTIEVNLPRIVGVCAKIMKLR 660
QY 657 KILEKVEAASGFTSEKOPBEFNILFHILRLRVEPLKIRSAQKQVDCYFYQIFMEKNE 716
DB 661 KILEKVEAASGFTSEKOPBEFNILFHILRLRVEPLKIRSAQKQVDCYFYQIFMEKNE 720
QY 717 KVCVPTIQOLLEWFSINSNLKFAEAPSLIIQMPRFKOPFKLPKIPPSLELNTDLED 776
DB 721 KVCVPTIQOLLEWFSINSNLKFAEAPSLIIQMPRFKOPFKLPKIPPSLELNTDLED 780
QY 777 TPRQCRICGLWAMECRECYDDPDISAGKIKQFCKTCTQVHLHPKRLNHNKYNPVS LPKD 836

DB 781 TPRQCRICGLWAMECRECYDDPDISAGKIKQFCKTCTQVHLHPKRLNHNKYNPVS LPKD 840
QY 837 LPDWDWRHGCIPQNNMELFAVLICIEISHVAVFYKYGKDSAWLFFDSDMADRGGQNGFNI 896
DB 841 LPDWDWRHGCIPQNNMELFAVLICIEISHVAVFYKYGKDSAWLFFDSDMADRGGQNGFNI 900
QY 897 PQVTPCPEVGEYVKMSLEDLHSLDSRRIOGCARRLLCDAYMCYQSTMSLYK 949
DB 901 PQVTPCPEVGEYVKMSLEDLHSLDSRRIOGCARRLLCDAYMCYQSTMSLYK 953

RESULT 2

ID ADQ95918 standard; protein; 953 AA.

XX AC ADQ95918;
XX DT 07-OCT-2004 (first entry)
XX T cell activation associated protein #48.

DE DE anti-allergic; antiarthritic; antiasthmatic; antidiabetic; anti-HIV;
KW antimicrobial; antirheumatic; immunosuppressive; neuroprotective;
KW gene therapy; T cell activation; diagnosis; autoimmune disease;
KW rheumatoid arthritis; asthma; multiple sclerosis; diabetes;
KW allergic disease; infectious disease; AIDS; chronic rejection; organ;
KW bone-marrow transplant.

XX OS Homo sapiens.

XX PN WO2004058805-A2.

XX PD 15-JUL-2004.

XX PF 25-DEC-2003; 2003WO-JP016715.

XX PR 26-DEC-2002; 2002JP-00376365.

XX PR 27-DEC-2002; 2002US-0436473P.

XX PR 25-APR-2003; 2003JP-00122113.

XX PR 28-APR-2003; 2003US-0465792P.

XX PR 21-OCT-2003; 2003JP-00360559.

XX PR 22-OCT-2003; 2003US-0512846P.

XX PA (ASAH-) ASAH KASEI PHARMA CORP.

XX PI Matsuda A, Yoneta S;

XX DR WPI; 2004-593134/57.

XX DR N-PSDB; ADQ95917.

XX PT New purified protein involved in T cell activation, useful for

XX PT diagnosing, preventing and/or treating acquired immunodeficiency

XX PT syndrome, autoimmune (e.g. rheumatoid arthritis, and diabetes), allergic

XX PT and infectious diseases.

XX PS Claim 1; SEQ ID NO 96; 2828pp; English.

XX CC The invention relates to purified proteins and genes encoding them, that

XX CC are involved in T cell activation (1) and has an amino acid deletion,

XX CC substitution or addition in the amino acid sequences. The methods and

XX CC compositions of the present invention are useful for the diagnosis,

XX CC prevention and/or treatment of autoimmune disease (rheumatoid arthritis,

XX CC asthma, multiple sclerosis and diabetes), allergic disease, infectious

XX CC disease, AIDS, and acute or chronic rejection at organ transplant or bone

XX CC marrow transplant. This sequence corresponds to a protein involved in T

XX CC cell activation.

SQ Sequence 953 AA;

Query Match

Best Local Similarity 99.0%; Score 4983; DB 8; Length 953;

Matches 948; Conservative 0; Mismatches 1; Indels 4; Gaps 4;

note

QY 1 MSSGLWSQEKVTSFYWEERIFVYLLLOECSTVDKOTQKLLKVPKGSIGQYIDRSVGHSHRI 60
Db 1 MSSGLWSQEKVTSFYWEERIFVYLLLOECSTVDKOTQKLLKVPKGSIGQYIDRSVGHSHRI 60
QY 61 PSAGKKNQIGLKILOPHAVLFVDE - DVVINEKFTTELLAITNCEERFSLFKNNRSL 119
Db 61 PSAGKKNQIGLKILOPHAVLFVDEKDVVEINEKFTTELLAITNCEERFSLFKNNRSL 120
QY 120 KGLQIDVGCVPVKQLRSGBEKFPGVVRFRGPLLAERTVSGIFPGVLLBEGRGQFTDGV 179
Db 121 KGLQIDVGCVPVKQLRSGBEKFPGVVRFRGPLLAERTVSGIFPGVLLBEGRGQFTDGV 180
QY 180 YQKQLFOCDEDCG - FVALDKLEIEDDDTALES DYAGPDTMQVELPLEINSRSLKG 238
Db 181 YQKQLFOCDEDCG FVALDKLEIEDDDTALES DYAGPDTMQVELPLEINSRSLKV 240
QY 239 GETIESGTIVFCVDLPKGSIGLVFGVDMNDPNTGNWDRFDGV - LCSFACVESTILLHN 297
Db 241 GETIESGTIVFCVDLPKGSIGLVFGVDMNDPNTGNWDRFDGVQVLCFACVESTILLHN 300
QY 298 DIIPESVTQERRPPKPLAFMRSGVDGKSSSHNPKATGSTDPGNR - RSELFYTLNGSSV 356
Db 301 DIIPESVTQERRPPKPLAFMRSGVDGKSSSHNPKATGSTDPGNRSELFTYTLNGSSV 360
QY 357 DSQOSKSNWYIIDEVAEDPAKSLTEISTDFDRSPPIOPPVPVNSLTENRPHSLPFSL 416
Db 361 DSQOSKSNWYIIDEVAEDPAKSLTEISTDFDRSPPIOPPVPVNSLTENRPHSLPFSL 420
QY 417 TQMPNTNGSIGHSPLSLSAQSWMEELNAPQESPPPLAMPNGSHGLEVGS LAEVKENP 476
Db 421 TQMPNTNGSIGHSPLSLSAQSWMEELNAPQESPPPLAMPNGSHGLEVGS LAEVKENP 480
QY 477 FYGVIRWIGQPLNEVLAGLEBECACGTDTGTRGTYFTYFCTALKKALFVKLKS CRPDS 536
Db 481 FYGVIRWIGQPLNEVLAGLEBECACGTDTGTRGTYFTYFCTALKKALFVKLKS CRPDS 540
QY 537 RFASLPQVSNQIERCNSLAFGGYLSVVEENTPPPKMEKEGLEIMICKKGIQGHYNSCYL 596
Db 541 RFASLPQVSNQIERCNSLAFGGYLSVVEENTPPPKMEKEGLEIMICKKGIQGHYNSCYL 600
QY 597 DSTLFCFAFSSVLTLLRPKEKNDVEYSETQELLRTTEIVNPLRIYGVVCATKIMKLR 656
Db 601 DSTLFCFAFSSVLTLLRPKEKNDVEYSETQELLRTTEIVNPLRIYGVVCATKIMKLR 660
QY 657 KILEKVEAASGFTSEKDPPEEFNLILFHHILRVEPLLKIRSAQKQVDCVYQIFMEKNE 716
Db 661 KILEKVEAASGFTSEKDPPEEFNLILFHHILRVEPLLKIRSAQKQVDCVYQIFMEKNE 720
QY 717 KVGVPITQOLLEWSFINSNLKFAEAPSCILIIQMPREGKDFKPKIFPSPLELNIITDLED 776
Db 721 KVGVPITQOLLEWSFINSNLKFAEAPSCILIIQMPREGKDFKPKIFPSPLELNIITDLED 780
QY 777 TPRQCRICGLIAMEYECREYDPPDISAGIKQPKCTCNTQVHLHPRKLNHKNYPSLPKD 836
Db 781 TPRQCRICGLIAMEYECREYDPPDISAGIKQPKCTCNTQVHLHPRKLNHKNYPSLPKD 840
QY 837 LPDWRHRCICPQNNELFAVICIETSHVYVAFVKYKDDSAWLFPSMADRDCGQNGFNI 896
Db 841 LPDWRHRCICPQNNELFAVICIETSHVYVAFVKYKDDSAWLFPSMADRDCGQNGFNI 900
QY 897 PQTTPCPEVGEYKMSLEDLHSLDSRRIQGCARRLLCDAVMCMYQPTMSLYK 949
Db 901 PQTTPCPEVGEYKMSLEDLHSLDSRRIQGCARRLLCDAVMCMYQPTMSLYK 953
RESULT 3
ADR99244
ID ADR99244 standard; protein; 960 AA.
XX
AC ADR99244;
XX
DT 02-DEC-2004 (first entry)

XX
DE DKFp586D1122, SEQ ID 250.
XX
KW Cytostatic; breast cancer; cancer; human; DKFp586D1122.
XX
OS Homo sapiens.
XX
PN WO2004078035-A2.
XX
PD 16-SEP-2004.
XX
PF 27-FEB-2004; 2004WO-US007268.
XX
PR 28-FEB-2003; 2003US-0450655P.
XX
PA (FARB) BAYER PHARM CORP.
XX
PI Eveleigh D, Bigwood D;
XX
XX WPI; 2004-653556/63.
DR N-PSDB; ADR99117.
DR
XX
XX
PT Diagnosing breast cancer comprises comparing the level of expression of
PT genes or gene products in a first biological sample taken from a patient
PT with that in a normal patient sample.
XX
PS Claim 3; SEQ ID NO 250; 53pp; English.
XX
CC The present invention relates to a method (M1) for diagnosing breast
CC cancer in a patient. The method comprises comparing the level of
CC expression of one or more genes or gene products in a biological sample
CC from the patient with that in a normal patient sample, where a difference
CC in the gene expression in the first sample compared to that in the second
CC sample is a diagnostic of the disease. Also claimed are: method (M2) for
CC distinguishing between normal and disease tissues; method (M3) for
CC monitoring the response of a breast cancer patient to treatment with an
CC anti-cancer agent; method (M4) for identifying a compound for treating
CC breast cancer; and an array for distinguishing between normal and disease
CC tissues comprising two or more probes corresponding to genes selected
CC from ADR98995-ADR99121 or comprising two or more polypeptides selected
CC from ADR99122-ADR99248. In M1 and M2 the genes are selected from ADR98995
CC -ADR99121 and the gene products are polypeptides selected from ADR99122-
CC ADR99248. M1 is useful for diagnosing breast cancer. M2 and the array are
CC useful for distinguishing between normal and disease tissue. M3 is useful
CC for monitoring the response of a breast cancer patient to treatment with
CC an anti-cancer agent. M4 is useful for identifying a compound for
CC treating breast cancer. Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 960 AA;

Query Match 99.0%; Score 4983; DB 8; Length 960;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 948; Conservative 0; Mismatches 1; Indels 4; Gaps 4;

QY 1 MSSGLWSQEKVTSFYWEERIFVYLLLOECSTVDKOTQKLLKVPKGSIGQYIDRSVGHSHRI 60
Db 8 MSSGLWSQEKVTSFYWEERIFVYLLLOECSTVDKOTQKLLKVPKGSIGQYIDRSVGHSHRI 67
QY 61 PSAGKKNQIGLKILOPHAVLFVDE - DVVINEKFTTELLAITNCEERFSLFKNNRSL 119
Db 68 PSAGKKNQIGLKILOPHAVLFVDEKDVVEINEKFTTELLAITNCEERFSLFKNNRSL 127
QY 120 KGLQIDVGCVPVKQLRSGBEKFPGVVRFRGPLLAERTVSGIFPGVLLBEGRGQFTDGV 179
Db 128 KGLQIDVGCVPVKQLRSGBEKFPGVVRFRGPLLAERTVSGIFPGVLLBEGRGQFTDGV 187
QY 180 YQKQLFOCDEDCG - FVALDKLEIEDDDTALES DYAGPDTMQVELPLEINSRSLKG 238
Db 188 YQKQLFOCDEDCG FVALDKLEIEDDDTALES DYAGPDTMQVELPLEINSRSLKV 247
QY 239 GETIESGTIVFCVDLPKGSIGLVFGVDMNDPNTGNWDRFDGV - LCSFACVESTILLHN 297

Db 248 GETIESGTIVFCVLPKESLGYFVGVDMDNPIGNWDRFDGVLQCSFACVESTILLHN 307
Qy 298 DIIPESVTQERRPPKLAFLMRGVDGKSSSHNKPATGSTSDPGNR-RSELEFYTLNGSSV 356
Db 308 DIIPESVTQERRPPKLAFLMRGVDGKSSSHNKPATGSTSDPGNRNRSLEFYTLNGSSV 367
Qy 357 DSQPSKSNWTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSLTNFRHSLPFSL 416
Db 368 DSQPSKSNWTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSLTNFRHSLPFSL 427
Qy 417 TKMPNTNGSIGHSPLSLSAQSVMEELNTPAQESPLAMPNGSHGLEVGSIAEVKENPP 476
Db 428 TKMPNTNGSIGHSPLSLSAQSVMEELNTPAQESPLAMPNGSHGLEVGSIAEVKENPP 487
Qy 477 FYGVIRWIGQPPGLNEVLAGELEDEACAGCTDGTFRGTRYFTCALKKALFVKLSCRPDS 536
Db 488 FYGVIRWIGQPPGLNEVLAGELEDEACAGCTDGTFRGTRYFTCALKKALFVKLSCRPDS 547
Qy 537 RFASLQPVSNQIERCNSLAFGGYLSSEWENTPPKMEKEGLEIMIGKKGIQGHVNSCYL 596
Db 548 RFASLQPVSNQIERCNSLAFGGYLSSEWENTPPKMEKEGLEIMIGKKGIQGHVNSCYL 607
Qy 597 DSTLFCCLFAPSSVLTLLRPKEKNDVEYYSETOELLRTTEINPLRIYGVYCATKIMKLR 656
Db 608 DSTLFCCLFAPSSVLTLLRPKEKNDVEYYSETOELLRTTEINPLRIYGVYCATKIMKLR 667
Qy 657 KILEKVEAASGTSREKOPPEEFNLIFHLRVEPLLKIRSAQKQVQDCVYFQIWEKNE 716
Db 668 KILEKVEAASGTSREKOPPEEFNLIFHLRVEPLLKIRSAQKQVQDCVYFQIWEKNE 727
Qy 717 KVGVTPIQQLLEWSFINSNLKFAEAPSCLIIQMPREGKDFKLFKIFPSLELNITDLED 776
Db 728 KVGVTPIQQLLEWSFINSNLKFAEAPSCLIIQMPREGKDFKLFKIFPSLELNITDLED 787
Qy 777 TPQCRICGLAMYECCREYDDPDISAGKIKQFCKTCNTQVHLHPKRLNHNKPNVSLPKD 836
Db 788 TPQCRICGLAMYECCREYDDPDISAGKIKQFCKTCNTQVHLHPKRLNHNKPNVSLPKD 847
Qy 837 LPDMDWRHGCIPQNMELFAVICIETSHYVAFVKYKQDSAMLFPDSMADRGQNGFNI 896
Db 848 LPDMDWRHGCIPQNMELFAVICIETSHYVAFVKYKQDSAMLFPDSMADRGQNGFNI 907
Qy 897 PQVTPCPEVGEYLNKWSLEDLHSLDSRIQGCARRLLCDAYMCYQSPMTSLYK 949
Db 908 PQVTPCPEVGEYLNKWSLEDLHSLDSRIQGCARRLLCDAYMCYQSPMTSLYK 960

RESULT 4

ADQ95916

ID ADQ95916 standard; protein; 953 AA.

XX AC ADQ95916;

XX DT 07-OCT-2004 (first entry)

XX DE T cell activation associated protein #47.

XX KW antiallergic; antiarthritic; antiaesthetic; antidiabetic; anti-HIV;
XX KW antimicrobial; antirheumatic; immunosuppressive; neuroprotective;
KW gene therapy; T cell activation; diagnosis; autoimmune disease;
KW rheumatoid arthritis; asthma; multiple sclerosis; diabetes;
KW allergic disease; infectious disease; AIDS; chronic rejection; organ;
KW bone-marrow transplant.

XX OS Homo sapiens.

XX PN WO2004058805-A2.

XX PD 15-JUL-2004.

XX PF 25-DEC-2003; 2003WO-JP016715.

XX

XX

XX

XX

XX

XX

XX

PR 26-DEC-2002; 2002JP-00376365.
PR 27-DEC-2002; 2002US-0436473P.
PR 25-APR-2003; 2003JP-00122113.
PR 28-APR-2003; 2003US-0465792P.
PR 21-OCT-2003; 2003JP-00360559.
PR 22-OCT-2003; 2003US-0512846P.

PA (ASAH-) ASahi Kasei Pharma Corp.

XX PI Matsuda A, Yoneta S;

XX WPI: 2004-593134/57.

DR N-PSDB; ADQ95915.

XX

PT New purified protein involved in T cell activation, useful for
PT diagnosing, preventing and/or treating acquired immunodeficiency
PT syndrome, autoimmune (e.g. rheumatoid arthritis, and diabetes), allergic
PT and infectious diseases.

PS Claim 1; SEQ ID NO 94; 2828pp; English.

XX CC The invention relates to purified proteins and genes encoding them, that
CC are involved in T cell activation (I) and has an amino acid deletion,
CC substitution or addition in the amino acid sequences. The methods and
CC compositions of the present invention are useful for the diagnosis,
CC prevention and/or treatment of autoimmune disease (rheumatoid arthritis,
CC asthma, multiple sclerosis and diabetes), allergic disease, infectious
CC disease, AIDS, and acute or chronic rejection at organ transplant or bone
CC marrow transplant. This sequence corresponds to a protein involved in T
CC cell activation.

XX SQ Sequence 953 AA;

Query Match 98.9%; Score 4980; DB 8; Length 953;

Best Local Similarity 99.4%; Pred. No. 0;

Matches 947; Conservative 1; Mismatches 1; Indels 4; Gaps 4;

Qy 1 MSSGLWSQEKVTSFYWEERIFYLLOECSTVDKQTKLLKVPKSGIQYIQDRSVGHSRI 60

Db 1 MSSGLWSQEKVTSFYWEERIFYLLOECSTVDKQTKLLKVPKSGIQYIQDRSVGHSRI 60

Qy 61 PSAKKNQIQLKILQPHAVLFVDE-DVVEINEKFTTELLAINTCERESLFPKRNRLS 119

Db 61 PSAKKNQIQLKILQPHAVLFVDEKVVVEINEKFTTELLAINTCERESLFPKRNRLS 120

Qy 120 KGLQIDVGCPCVKQLRSGEKEKPGVVRFRGPLLAERTVSGIFGVLELLEGRGOGFTDGV 179

Db 121 KGLQIDVGCPCVKQLRSGEKEKPGVVRFRGPLLAERTVSGIFGVLELLEGRGOGFTDGV 180

Qy 180 YQKQLFQDCDECG-FVALDKLELIEDDDTALSDYAGPGDTMQVELPPLLEINSRSLKG 238

Db 181 YQKQLFQDCDECGFVALDKLELIEDDDTALSDYAGPGDTMQVELPPLLEINSRSLKV 240

Qy 239 GETIESGTIVFCVLPKESLGYFVGVDMDNPIGNWDRFDGVLQCSFACVESTILLHN 297

Db 241 GETIESGTIVFCVLPKESLGYFVGVDMDNPIGNWDRFDGVLQCSFACVESTILLHN 300

Qy 298 DIIPESVTQERRPPKLAFLMRGVDGKSSSHNKPATGSTSDPGNR-RSELEFYTLNGSSV 356

Db 301 DIIPESVTQERRPPKLAFLMRGVDGKSSSHNKPATGSTSDPGNRNRSLEFYTLNGSSV 360

Qy 357 DSQPSKSNWTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSLTNFRHSLPFSL 416

Db 361 DSQPSKSNWTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSLTNFRHSLPFSL 420

Qy 417 TKMPNTNGSIGHSPLSLSAQSVMEELNTPAQESPLAMPNGSHGLEVGSIAEVKENPP 476

Db 421 TKMPNTNGSIGHSPLSLSAQSVMEELNTPAQESPLAMPNGSHGLEVGSIAEVKENPP 480

Qy 477 FYGVIRWIGQPPGLNEVLAGELEDEACAGCTDGTFRGTRYFTCALKKALFVKLSCRPDS 536

Db 481 FYGVIRWIGQPPGLNEVLAGELEDEACAGCTDGTFRGTRYFTCALKKALFVKLSCRPDS 540

3
data

QY 537 REASLOPVSNQIERNCSLAFGGYLSVVEENTPPKMEKGLIEMICKKGIQGHVNSCYL 596
DB 541 RFASLOPVSNQIERNCSLAFGGYLSVVEENTPPKMEKGLIEMICKKGIQGHVNSCYL 600
QY 597 DSTLFCFAFSSVLDIVLLRPKEKNDVEYYSQELLRTEIVNPLRIYGVYCATKIMKLR 656
DB 601 DSTLFCFAFSSVLDIVLLRPKEKNDVEYYSQELLRTEIVNPLRIYGVYCATKIMKLR 660
QY 657 KILEKVEAASGTSSEKDPPEEFNLFHILRVEPLLKIRSAQKQVQDCYFYQIFMEKNE 716
DB 661 KILEKVEAASGTSSEKDPPEEFNLFHILRVEPLLKIRSAQKQVQDCYFYQIFMEKNE 720
QY 717 KVGVPVPTIOQLLEWSFINSNLKFAEAPSLIIQMPREGKDFPKFKIPPSLELNITDLE 776
DB 721 KVGVPVPTIOQLLEWSFINSNLKFAEAPSLIIQMPREGKDFPKFKIPPSLELNITDLE 780
QY 777 TPRQCRIICGLAMYECCRECYDDPDISAGKIKQFCKTCNTQVHLHPKRLNHNKYNPVSILPKD 836
DB 781 TPRQCRIICGLAMYECCRECYDDPDISAGKIKQFCKTCNTQVHLHPKRLNHNKYNPVSILPKD 840
QY 837 LPDWDWRHGCIPCONMELFAVLCIETSHYVAFVKYKDDSAWLFFDSDMADRDGGQNGFNI 896
DB 841 LPDWDWRHGCIPCONMELFAVLCIETSHYVAFVKYKDDSAWLFFDSDMADRDGGQNGFNI 900
QY 897 PQTPTCPPEVGEYLKMSLEDLHLSDSRRIQGCARRLLCDAYMCMYQSPPTMSLYK 949
DB 901 PQTPTCPPEVGEYLKMSLEDLHLSDSRRIQGCARRLLCDAYMCMYQSPPTMSLYK 953

RESULT 5

ABB82783
ID ABB82783 standard; protein; 956 AA.

AC ABB82783;
DT 18-MAR-2003 (first entry)
XX Human CYLD polypeptide.

DE CD40; NF-kappaB essential modulator; NEMO; CYLD; immunosuppressive;
KW antiinflammatory; cytostatic; gene therapy; human.
XX Homo sapiens.

OS WO200292761-A2.
XX 21-NOV-2002.

PD 07-MAY-2002; 2002WO-US014570.
XX .08-MAY-2001; 2001US-00851673.

PR (IMV) IMMUNEX CORP.
XX Derry JMJ, Fanslow WC, Dougall WC;
XX WPI; 2003-120669/11.
XX N-PSDB; ABV75394.

DR Identifying compounds that alter one or more biological activities of
XX CD40 by modulating the binding of NEMO and CYLD, useful for treating
XX disorders of the immune system, and inflammatory and cancer diseases.

PS Claim 1; Page 38-44; 48pp; English.

CC The invention relates to methods of identifying compounds that alter one/
CC more biological activities of CD40. One method involves screening for
CC molecules that modulate the binding of NF-kappaB essential modulator
CC (NEMO) and CYLD. The methods and compositions of the invention of
CC determining compounds that agonize or antagonize a CD40 signaling
CC activity, are useful for the further definition of CD40-mediated
CC signaling pathways, and for manipulation of CD40-mediated cellular
CC responses. They also provide therapeutic agents for treating disorders of

CC the immune system, and inflammatory and cancer diseases. The present
CC sequence represents the human CYLD polypeptide
XX
SQ Sequence 956 AA;

Query Match 98.8%; Score 4971.5; DB 6; Length 956;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 948; Conservative 0; Mismatches 1; Indels 7; Gaps 5;

QY 1 MSSGLMSQEKVTSFYWEERIFYLLQECVSTDKQTKLLKVPKGSIGQYIODRSVGHRSI 60
DB 1 MSSGLMSQEKVTSFYWEERIFYLLQECVSTDKQTKLLKVPKGSIGQYIODRSVGHRSI 60
QY 61 PSAGKKNQIKILQOPHAFVDE--DVVEINEKFTTELLAITNCBERFSLFKNRNLS 119
DB 61 PSAGKKNQIKILQOPHAFVDEKDVVEINEKFTTELLAITNCBERFSLFKNRNLS 120
QY 120 KGIQIDVGCVPKQVLRSGBEKFPQVFRFRPLAERTVSGIFFGVVLEELGEGQFTDGV 179
DB 121 KGIQIDVGCVPKQVLRSGBEKFPQVFRFRPLAERTVSGIFFGVVLEELGEGQFTDGV 180
QY 180 YQKQLFQCDDECG-FVALDKLEIEDDDTALSDYAGPGDTMVELPPLPINSRVSILKG 238
DB 181 YQKQLFQCDDECGVFVALDKLEIEDDDTALSDYAGPGDTMVELPPLPINSRVSILKV 240
QY 239 GETIESGTVIFCDVLPKESLGYFVGVDMPNIPGNWDGRFDGV-LCSFACVESTILLHN 297
DB 241 GETIESGTVIFCDVLPKESLGYFVGVDMPNIPGNWDGRFDGVLCSPACVESTILLHN 300
QY 298 DIIP---ESVTOERRPPKLAFMGRGVGDKGSSSHNKPKATGSTSDPGR-RSELYFTLNG 353
DB 301 DIIPALSESIVTOERRPPKLAFMGRGVGDKGSSSHNKPKATGSTSDPGRNRSELYFTLNG 360
QY 354 SSVDSQPSQSKNTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSLTTERNFHSLP 413
DB 361 SSVDSQPSQSKNTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSLTTERNFHSLP 420
QY 414 FSLTKMPTNGSIGHSPLSLSAQSVMBELNAPVQESPPPLAMPFGNSHGLEVGSIAEYKE 473
DB 421 FSLTKMPTNGSIGHSPLSLSAQSVMBELNAPVQESPPPLAMPFGNSHGLEVGSIAEYKE 480
QY 474 NPPFYGVIRWIGOPGLNEVLAGELEDEACAGTGTGFRGTRYFTCALKKALFKVLKSCR 533
DB 481 NPPFYGVIRWIGOPGLNEVLAGELEDEACAGTGTGFRGTRYFTCALKKALFKVLKSCR 540
QY 534 PDSRFASLOPVSNQIERCNSLAFGGYLSVVEENTPPKMEKGLIEMICKKGIQGHVNS 593
DB 541 PDSRFASLOPVSNQIERCNSLAFGGYLSVVEENTPPKMEKGLIEMICKKGIQGHVNS 600
QY 594 CYLDSTLFCFAFSSVLDIVLLRPKEKNDVEYYSQELLRTEIVNPLRIYGVYCATKIM 653
DB 601 CYLDSTLFCFAFSSVLDIVLLRPKEKNDVEYYSQELLRTEIVNPLRIYGVYCATKIM 660
QY 654 KLRKILEKVEAASGTSSEKDPPEEFNLFHILRVEPLLKIRSAQKQVQDCYFYQIFME 713
DB 661 KLRKILEKVEAASGTSSEKDPPEEFNLFHILRVEPLLKIRSAQKQVQDCYFYQIFME 720
QY 714 KNEKVGVPVPTIOQLLEWSFINSNLKFAEAPSLIIQMPREGKDFPKFKIPPSLELNITDL 773
DB 721 KNEKVGVPVPTIOQLLEWSFINSNLKFAEAPSLIIQMPREGKDFPKFKIPPSLELNITDL 780
QY 774 LEDTPQCRIICGLAMYECCRECYDDPDISAGKIKQFCKTCNTQVHLHPKRLNHNKYNPVS 833
DB 781 LEDTPQCRIICGLAMYECCRECYDDPDISAGKIKQFCKTCNTQVHLHPKRLNHNKYNPVS 840
QY 834 PKDLPDWDWRHGCIPCONMELFAVLCIETSHYVAFVKYKDDSAWLFFDSDMADRDGGQNG 893
DB 841 PKDLPDWDWRHGCIPCONMELFAVLCIETSHYVAFVKYKDDSAWLFFDSDMADRDGGQNG 900
QY 894 FNIPTPTCPPEVGEYLKMSLEDLHLSDSRRIQGCARRLLCDAYMCMYQSPPTMSLYK 949
DB 901 FNIPTPTCPPEVGEYLKMSLEDLHLSDSRRIQGCARRLLCDAYMCMYQSPPTMSLYK 956

RESULT 6

ADRI14489
 ID ADRI14489 standard; protein; 956 AA.
 AC ADRI14489;
 XX
 DT 21-OCT-2004 (first entry)
 XX
 DE Human NF-kappaB pathway-associated protein SeqID490.
 XX
 KW NF-kappaB pathway; antiinflammatory; cytostatic; hepatotropic; virucide;
 KW antiarthritic; antirheumatic; gastrointestinal-gen; antiasthmatic;
 KW antiarteriosclerotic; immunomodulator; cerebroprotective; vasotropic;
 KW immunosuppressive; vulnary; gene therapy; immune disorder;
 KW inflammatory disorder; NF-kappaB regulation; cancer; aberrant apoptosis;
 KW hepatic disorder; Hodgkin's lymphoma; haematopoietic tumour;
 KW hyper-IgM syndrome; hypohidrotic ectodermal dysplasia;
 KW X-linked anhidrotic ectodermal dysplasia; immunodeficiency;
 KW viral infection; HIV-1; HTLV-1; hepatitis B; hepatitis C; EBV; influenza;
 KW viral replication; host cell survival; evasion of immune response;
 KW rheumatoid arthritis; inflammatory bowel disease; colitis; asthma;
 KW atherosclerosis; cachexia; euthyroid sick syndrome; stroke; EAE;
 KW autoimmune disorder; hyper immune activity;
 KW aberrant acute phase response; hypercongenital condition; birth defect;
 KW necrotic lesion; wound; organ transplant rejection;
 KW aberrant signal transduction; proliferating disorder; cancer;
 KW HIV propagation; human.
 XX
 OS Homo sapiens.
 XX
 PN WO2004065577-A2.
 XX
 PD 05-AUG-2004.
 XX
 PF 13-JAN-2004; 2004WO-US000798.
 XX
 PF 14-JAN-2003; 2003US-0440068P.
 PR 12-MAY-2003; 2003US-0469757P.
 XX
 PA (BRIM) BRISTOL-MYERS SQUIBB CO.
 XX
 PI Nadler SG, Neubauer MG, Feder JN, Carman J;
 XX
 DR WPI; 2004-562168/54.
 DR N-PSDB; ADRI14488.
 XX
 XX New isolated polynucleotides and polypeptides associated with NF-kappaB
 PT pathway, useful for diagnosing, treating, or preventing disorders or
 PT diseases associated with NF-kappaB pathway.
 XX
 PS Claim 6; SEQ ID NO 490; 237pp; English.
 XX
 CC This invention relates to the novel association of protein sequences (and
 CC the genes which encode them) to the NF-kappaB pathway. The invention may
 CC be useful for the production of compounds with an antiinflammatory,
 CC cytostatic, hepatotropic, virucide, antiarthritic, antirheumatic,
 CC gastrointestinal-gen, antiasthmatic, antiarteriosclerotic,
 CC immunomodulator, cerebroprotective, vasotropic, immunosuppressive or
 CC vulnary activity or for gene therapy. The proteins and nucleotides are
 CC useful for diagnosing, preventing, treating, or ameliorating conditions
 CC or diseases associated with the NF-kappaB pathway. The condition is an
 CC immune disorder, an inflammatory disorder, an inflammatory disorder
 CC related to aberrant NF-kappaB regulation, cancer, aberrant apoptosis,
 CC hepatic disorders, Hodgkin's lymphomas, haematopoietic tumours, hyper-IgM
 CC syndromes, hypohidrotic ectodermal dysplasia, X-linked anhidrotic
 CC ectodermal dysplasia, immunodeficiency, viral infections, host cell
 CC hepatitis B, hepatitis C, EBV, influenza, viral replication, host cell
 CC survival, evasion of immune responses, rheumatoid arthritis, inflammatory
 CC bowel disease, colitis, asthma, atherosclerosis, cachexia, euthyroid sick
 CC syndrome, stroke, EAE, autoimmune disorders, disorders related to hyper
 CC immune activity, disorders related to aberrant acute phase responses,
 CC hypercongenital conditions, birth defects, necrotic lesions, wounds,

CC organ transplant rejection, conditions related to organ transplant
 CC rejection, disorders related to aberrant signal transduction, infected
 CC proliferating disorders, cancers and HIV propagation in cells which
 CC is subject to the novel association with the NF-kappaB pathway of the
 CC invention. Note: This sequence does not appear in the specification but
 CC was obtained by the indexer from Genbank.
 XX
 SQ Sequence 956 AA;
 Query Match 98.8%; Score 4971.5; DB 8; Length 956;
 Best Local Similarity 99.2%; Pred. No. 0;
 Matches 948; Conservative 0; Mismatches 1; Indels 7; Gaps 5;
 QY 1 MSSGLWSQEKVTSPYWEERIFVLLQLQECSTVDKOTQKLLKVPKSGISQYIQDRSVGHSRI 60
 DB 1 MSSGLWSQEKVTSPYWEERIFVLLQLQECSTVDKOTQKLLKVPKSGISQYIQDRSVGHSRI 60
 QY 61 PSAKGGKNOIGLILQPHAVLFDVE-DVVEINEKFTTELLAITNCBERSLFKNRNRLS 119
 DB 61 PSAKGGKNOIGLILQPHAVLFDVEKDVVEINEKFTTELLAITNCBERSLFKNRNRLS 120
 QY 120 KGLQIDVGCPCVKQLRSGBEKPFGVVRFRGPLLAERTVSGIFFGVELLEBGRGQFTDGV 179
 DB 121 KGLQIDVGCPCVKQLRSGBEKPFGVVRFRGPLLAERTVSGIFFGVELLEBGRGQFTDGV 180
 QY 180 YQKQLFQCEDECG-FVALDKLELIEDDDTALDES DYAGPGDTMQVELPPLPPEINRSVLKG 238
 DB 181 YQKQLFQCEDECGFVALDKLELIEDDDTALDES DYAGPGDTMQVELPPLPPEINRSVLKV 240
 QY 239 GETIESGTIVFCDLPGKSLGYFVGVDMDNPIGNWDGRPDGV-LCSFACVESTILLHIN 297
 DB 241 GETIESGTIVFCDLPGKSLGYFVGVDMDNPIGNWDGRPDGV-LCSFACVESTILLHIN 300
 QY 298 DIIP---ESVTOERRPPKLAFAKSRGVDKSGSSSHNPKATGSTDGCR-RSELFYTLNG 353
 DB 301 DIIPALSESVTOERRPPKLAFAKSRGVDKSGSSSHNPKATGSTDGCRNRSELUFYTLNG 360
 QY 354 SSVDSQPSQSKNTWYIDEVAEDPAKSLTEISTDFDRSSPPLPAPPVNSLTENRHSPL 413
 DB 361 SSVDSQPSQSKNTWYIDEVAEDPAKSLTEISTDFDRSSPPLPAPPVNSLTENRHSPL 420
 QY 414 FSLTKMPTNGSIGHSPLSLSAQSVMEELNTPVQSSPPLAMPNGSHGLEVGS LAEVE 473
 DB 421 FSLTKMPTNGSIGHSPLSLSAQSVMEELNTPVQSSPPLAMPNGSHGLEVGS LAEVE 480
 QY 474 NPPFYGIWIRGIPGPGINEVLAGELEDECACTDGTFRGTRYFTCALKKALFVKLKSCR 533
 DB 481 NPPFYGIWIRGIPGPGINEVLAGELEDECACTDGTFRGTRYFTCALKKALFVKLKSCR 540
 QY 534 PDSRFASLOPVSNQIERCNSLAFGGYLSVEVENTPPKMEKGLIEMIGKKGIQGHYNS 593
 DB 541 PDSRFASLOPVSNQIERCNSLAFGGYLSVEVENTPPKMEKGLIEMIGKKGIQGHYNS 600
 QY 594 CYLDSTLFCFLAFSSVLDVTLVLRPKNDVEYSETQELLRTIENPLRIYGVYCATKIM 653
 DB 601 CYLDSTLFCFLAFSSVLDVTLVLRPKNDVEYSETQELLRTIENPLRIYGVYCATKIM 660
 QY 654 KLRKILKEVEAASGFTSEKDPPEFLNPLFHILRVESPLLKIRSAGQKVDQCFYQIFME 713
 DB 661 KLRKILKEVEAASGFTSEKDPPEFLNPLFHILRVESPLLKIRSAGQKVDQCFYQIFME 720
 QY 714 KNEKVGPTIQQLLEWSFINSNLKFAEAPSCLLIQMPFGCKDFKLFKKIPPSLELNTDL 773
 DB 721 KNEKVGPTIQQLLEWSFINSNLKFAEAPSCLLIQMPFGCKDFKLFKKIPPSLELNTDL 780
 QY 774 LEDTPRQCRI CGGLAMVECRECYDDPDISAGKIKQFCKTCTNTQVHLHPKRLNHYNPVSL 833
 DB 781 LEDTPRQCRI CGGLAMVECRECYDDPDISAGKIKQFCKTCTNTQVHLHPKRLNHYNPVSL 840
 QY 834 PKDLPDWDRHGCIPQNMELFAVLCTIETSHYVAFVKYKDDSAWLPFDSMADRGQNG 893
 DB 841 PKDLPDWDRHGCIPQNMELFAVLCTIETSHYVAFVKYKDDSAWLPFDSMADRGQNG 900

QY 894 FNIPQVTPCEVGEYLKMSLEDLHSLDSRRIOGCARRLLCDAYMCMYQSPMTSLYK 949
 DB 901 FNIPQVTPCEVGEYLKMSLEDLHSLDSRRIOGCARRLLCDAYMCMYQSPMTSLYK 956

RESULT 7

ADQ95920
 ID ADQ95920 standard; protein; 956 AA.

AC ADQ95920;

DT 07-OCT-2004 (first entry)

DE T cell activation associated protein #49.

XX anti-allergic; antiarthritic; antiaesthatic; antidiabetic; anti-HIV;
 KW antimicrobial; antirheumatic; immunosuppressive; neuroprotective;
 KW gene therapy; T cell activation; diagnosis; autoimmune disease;
 KW rheumatoid arthritis; asthma; multiple sclerosis; diabetes;
 KW allergic disease; infectious disease; AIDS; chronic rejection; organ;
 KW bone-marrow transplant.

XX Homo sapiens.

XX WO2004058805-A2.

PN 15-JUL-2004.

XX 25-DEC-2003; 2003WO-JP016715.

XX 26-DEC-2002; 2002JP-00376365.

PR 27-DEC-2002; 2002US-0436473P.

PR 25-APR-2003; 2003JP-00122113.

PR 28-APR-2003; 2003US-0465792P.

PR 21-OCT-2003; 2003JP-00360559.

PR 22-OCT-2003; 2003US-0512846P.

XX (ASAH-) ASahi Kasei Pharma Corp.

PI Matsuda A, Yoneta S;

XX WPI; 2004-593134/57.

DR N-PSDB; ADQ95919.

XX New purified protein involved in T cell activation, useful for

PT diagnosing, preventing and/or treating acquired immunodeficiency

PT syndrome, autoimmune (e.g. rheumatoid arthritis, and diabetes), allergic

PT and infectious diseases.

XX Claim 1; SEQ ID NO 98; 2828pp; English.

XX The invention relates to purified proteins and genes encoding them, that

CC are involved in T cell activation (I) and has an amino acid deletion,

CC substitution or addition in the amino acid sequences. The methods and

CC compositions of the present invention are useful for the diagnosis,

CC prevention and/or treatment of autoimmune disease (rheumatoid arthritis,

CC asthma, multiple sclerosis and diabetes), allergic disease, infectious

CC disease, AIDS, and acute or chronic rejection at organ transplant or bone

CC marrow transplant. This sequence corresponds to a protein involved in T

XX cell activation.

XX Sequence 956 AA;

DB 61 PSAGKGNQIGLKILEQPHAVLFVDEKDVVEINEKFTELLALATNCERSLFPNRRLS 120
 QY 120 KGLQIDVGCPCVKQLRSGBEEKFPGVVRFRGPLAAERTVSGIFFGVVLEEGRGQGTDCV 179
 DB 121 KGLQIDVGCPCVKQLRSGBEEKFPGVVRFRGPLAAERTVSGIFFGVVLEEGRGQGTDCV 180
 QY 180 YQKQLFOCEDDCG-FVALDKLELIEDDDTALESDDYAGDGTMOVELPPELINSRVS LG 238
 DB 181 YQKQLFOCEDDCGCVFVALDKLELIEDDDTALESDDYAGDGTMOVELPPELINSRVS LGV 240
 QY 239 GETIESGTVIFCDVLPKESLG YFVGVDMDNPNINMDGRFDGV-LCSFACVESTILLHIN 297
 DB 241 GETIESGTVIFCDVLPKESLG YFVGVDMDNPNINMDGRFDGV-LCSFACVESTILLHIN 300
 QY 298 DIIP---ESVTQERRPPKLA FMSRGVGDKGSSSHNKP KATGSTSDPQNR-RSELYFYTLNG 353
 DB 301 DIIPALSESVTQERRPPKLA FMSRGVGDKGSSSHNKP KATGSTSDPQNRSELFYTLNG 360
 QY 354 SSVDSQPSKSNKTWYIDEVAEDPAKSLTETISDFDRSSPPLQPPPNVSLTTERFHSPL 413
 DB 361 SSVDSQPSKSNKTWYIDEVAEDPAKSLTETISDFDRSSPPLQPPPNVSLTTERFHSPL 420
 QY 414 FSLTKMPTNGSIGHSPLSLSAQSVMEELNTAPVQESSPPLAMPNGSHGLEVGSLAEVKE 473
 DB 421 FSLTKMPTNGSIGHSPLSLSAQSVMEELNTAPVQESSPPLAMPNGSHGLEVGSLAEVKE 480
 QY 474 NPPFYGVIRWIGQPPGLNEVLAGELEDEACAGCTDGTFRGTRYFTCAKKAFLVKLXSCR 533
 DB 481 NPPFYGVIRWIGQPPGLNEVLAGELEDEACAGCTDGTFRGTRYFTCAKKAFLVKLXSCR 540
 QY 534 PDSRFASLOPVSNQIERCNSLAFGGYLSVVEENTPPKMEKEGLEIMIGKKKGIGHYNS 593
 DB 541 PDSRFASLOPVSNQIERCNSLAFGGYLSVVEENTPPKMEKEGLEIMIGKKKGIGHYNS 600
 QY 594 CYLDSTLFCLEAFSSVLDTVLLRPKKNVVEYSETQELLRTETVNPRLIYGVYCATKIM 653
 DB 601 CYLDSTLFCLEAFSSVLDTVLLRPKKNVVEYSETQELLRTETVNPRLIYGVYCATKIM 660
 QY 654 KLRKILEKVEAASGFTSEKDPPEFLNLPFHILRVEPLLKIRSAQGVQDCYFYQIFME 713
 DB 661 KLRKILEKVEAASGFTSEKDPPEFLNLPFHILRVEPLLKIRSAQGVQDCYFYQIFME 720
 QY 714 KNEKVGPTIQQLLEWSFINSNLKFAEAPSCLIIQMPRFGKDFKLPKIPPSLEINITDL 773
 DB 721 KNEKVGPTIQQLLEWSFINSNLKFAEAPSCLIIQMPRFGKDFKLPKIPPSLEINITDL 780
 QY 774 LEDTPRQCRICGGLAMVECRECYDDPDISAGKIKQFCKTQNTQVHLHPKRLNHNKYNPVS 833
 DB 781 LEDTPRQCRICGGLAMVECRECYDDPDISAGKIKQFCKTQNTQVHLHPKRLNHNKYNPVS 840
 QY 834 PKDLPDMDWRHGCIPQNMELFAVLCIETSHYVAFVKYKDDSAWLFDDSMADRDGGONG 893
 DB 841 PKDLPDMDWRHGCIPQNMELFAVLCIETSHYVAFVKYKDDSAWLFDDSMADRDGGONG 900
 QY 894 FNIPQVTPCEVGEYLKMSLEDLHSLDSRRIOGCARRLLCDAYMCMYQSPMTSLYK 949
 DB 901 FNIPQVTPCEVGEYLKMSLEDLHSLDSRRIOGCARRLLCDAYMCMYQSPMTSLYK 956

RESULT 8

AAU23747

ID AAU23747 standard; protein; 739 AA.

XX AAU23747;

XX DT 17-DEC-2001 (first entry)

XX Novel human enzyme polypeptide #833.

XX DE

XX Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;

KW ligase; hyperproliferative disorder; immunodeficiency disorder;

KW autoimmune disorder; neurological disorder; metabolic disorder;

KW inflammatory disorder; cardiovascular disorder; reproductive disorder;
KW blood-related disorder; infectious disorder; cytostatic; anti arthritic;
XX nephrotropic; anticoagulant.

OS Homo sapiens.

XX WO200155301-A2.

FN 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US001239.

XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.

PR 14-AUG-2000; 2000US-0225759P.

PR 18-AUG-2000; 2000US-0226279P.

PR 22-AUG-2000; 2000US-0226681P.

PR 22-AUG-2000; 2000US-0226688P.

PR 22-AUG-2000; 2000US-0227182P.

PR 23-AUG-2000; 2000US-0227009P.

PR 30-AUG-2000; 2000US-0228924P.

PR 01-SEP-2000; 2000US-0229287P.

PR 01-SEP-2000; 2000US-0229343P.

PR 01-SEP-2000; 2000US-0229344P.

PR 01-SEP-2000; 2000US-0229345P.

PR 05-SEP-2000; 2000US-0229509P.

PR 05-SEP-2000; 2000US-0229513P.

PR 06-SEP-2000; 2000US-0230437P.

PR 06-SEP-2000; 2000US-0230438P.

PR 08-SEP-2000; 2000US-0231242P.

PR 08-SEP-2000; 2000US-0231243P.

PR 08-SEP-2000; 2000US-0231244P.

PR 08-SEP-2000; 2000US-0231413P.

PR 08-SEP-2000; 2000US-0231414P.

PR 08-SEP-2000; 2000US-0232080P.

PR 08-SEP-2000; 2000US-0232081P.

PR 12-SEP-2000; 2000US-0231968P.

PR 14-SEP-2000; 2000US-0232397P.

PR 14-SEP-2000; 2000US-0232398P.

PR 14-SEP-2000; 2000US-0232399P.

PR 14-SEP-2000; 2000US-0232401P.

PR 14-SEP-2000; 2000US-0233063P.

PR 14-SEP-2000; 2000US-0233064P.

PR 14-SEP-2000; 2000US-0233065P.

PR 21-SEP-2000; 2000US-0234223P.

PR 21-SEP-2000; 2000US-0234274P.

PR 25-SEP-2000; 2000US-0234997P.

PR 25-SEP-2000; 2000US-0234998P.

PR 26-SEP-2000; 2000US-0235484P.

PR 27-SEP-2000; 2000US-0235834P.

PR 27-SEP-2000; 2000US-0235836P.

PR 29-SEP-2000; 2000US-0236327P.

PR 29-SEP-2000; 2000US-0236367P.

PR 29-SEP-2000; 2000US-0236368P.

PR 29-SEP-2000; 2000US-0236369P.

PR 29-SEP-2000; 2000US-0236370P.

PR 02-OCT-2000; 2000US-0236802P.

PR 02-OCT-2000; 2000US-0237037P.

PR 02-OCT-2000; 2000US-0237038P.

PR 02-OCT-2000; 2000US-0237039P.

PR 13-OCT-2000; 2000US-0237040P.

PR 13-OCT-2000; 2000US-0239935P.

PR 13-OCT-2000; 2000US-0239937P.

PR 20-OCT-2000; 2000US-0240960P.

PR 20-OCT-2000; 2000US-0241221P.

PR 20-OCT-2000; 2000US-0241785P.

PR 20-OCT-2000; 2000US-0241786P.

PR 20-OCT-2000; 2000US-0241787P.

PR 20-OCT-2000; 2000US-0241808P.

PR 20-OCT-2000; 2000US-0241809P.

PR 20-OCT-2000; 2000US-0241826P.

PR 01-NOV-2000; 2000US-0244617P.

PR 08-NOV-2000; 2000US-0246474P.

PR 08-NOV-2000; 2000US-0246475P.

PR 08-NOV-2000; 2000US-0246476P.

PR 08-NOV-2000; 2000US-0246477P.

PR 08-NOV-2000; 2000US-0246478P.

PR 08-NOV-2000; 2000US-0246523P.

PR 08-NOV-2000; 2000US-0246524P.

PR 08-NOV-2000; 2000US-0246525P.

PR 08-NOV-2000; 2000US-0246526P.

PR 08-NOV-2000; 2000US-0246527P.

PR 08-NOV-2000; 2000US-0246528P.

PR 08-NOV-2000; 2000US-0246532P.

PR 08-NOV-2000; 2000US-0246609P.

PR 08-NOV-2000; 2000US-0246610P.

PR 08-NOV-2000; 2000US-0246611P.

PR 08-NOV-2000; 2000US-0246613P.

PR 17-NOV-2000; 2000US-0249207P.

PR 17-NOV-2000; 2000US-0249208P.

PR 17-NOV-2000; 2000US-0249209P.

PR 17-NOV-2000; 2000US-0249210P.

PR 17-NOV-2000; 2000US-0249211P.

PR 17-NOV-2000; 2000US-0249212P.

PR 17-NOV-2000; 2000US-0249213P.

PR 17-NOV-2000; 2000US-0249214P.

PR 17-NOV-2000; 2000US-0249215P.

PR 17-NOV-2000; 2000US-0249216P.

PR 17-NOV-2000; 2000US-0249217P.

PR 17-NOV-2000; 2000US-0249218P.

PR 17-NOV-2000; 2000US-0249244P.

PR 17-NOV-2000; 2000US-0249245P.

PR 17-NOV-2000; 2000US-0249264P.

PR 17-NOV-2000; 2000US-0249265P.

PR 17-NOV-2000; 2000US-0249297P.

PR 17-NOV-2000; 2000US-0249299P.

PR 01-DEC-2000; 2000US-0250160P.

PR 01-DEC-2000; 2000US-0250391P.

PR 05-DEC-2000; 2000US-0251030P.

PR 05-DEC-2000; 2000US-0251988P.

PR 05-DEC-2000; 2000US-0256719P.

PR 06-DEC-2000; 2000US-0251479P.

PR 08-DEC-2000; 2000US-0251856P.

PR 08-DEC-2000; 2000US-0251868P.

PR 08-DEC-2000; 2000US-0251869P.

PR 08-DEC-2000; 2000US-0251989P.

PR 08-DEC-2000; 2000US-0251990P.

PR	11-DEC-2000; 2000US-0254097P.	QY	661	KVEAASGFTSEKDPPEFLNLPFHLLRVPEPLKIRSAGQVQDCYFYQIFMEKNEKVG	720
PR	05-JAN-2001; 2001US-0259678P.	Db	451	KVEAASGFTSEKDPPEFLNLPFHLLRVPEPLKIRSAGQVQDCYFYQIFMEKNEKVG	510
PA	(HUMA-) HUMAN GENOME SCI INC.	QY	721	PTIQQLLEWSFINSNLKFAEAPSCLIIQMPFRGKDFKLFKKI PPSLELNTDLEDTPRQ	780
XX	Rosen CA, Barash SC, Ruben SM;	Db	511	PTIQQLLEWSFINSNLKFAEAPSCLIIQMPFRGKDFKLFKKI PPSLELNTDLEDTPRQ	570
XX	WPI, 2001-465566/50.	QY	781	CRICGGLAMVECRECYDDPDISAGIKQFCKTQNTQVHLHPKRLNHKYNPVSPLPKDLPDW	840
DR	N-PSDB; AAS41617.	Db	571	CRICGGLAMVECRECYDDPDISAGIKQFCKTQNTQVHLHPKRLNHKYNPVSPLPKDLPDW	630
PT	Novel polypeptides and polynucleotides useful for diagnosing, preventing,	QY	841	DWRHGCIPQCNMELFAVLCTIETSHYVAFVKYKDDSAWLPFDSMADRDGGONGNIPOVT	900
PT	treating neural, immune system, muscular, reproductive, pulmonary,	Db	631	DWRHGCIPQCNMELFAVLCTIETSHYVAFVKYKDDSAWLPFDSMADRDGGONGNIPOVT	690
PT	cardiovascular, renal, proliferative disorders and cancerous diseases.	QY	901	PCPEVGEYLKMSLEDLSRRIOGCARRLLCDAYCMYQSPMTSLYK	949
XX	Claim 11; SEQ ID NO 1743; 1180pp; English.	Db	691	PCPEVGEYLKMSLEDLSRRIOGCARRLLCDAYCMYQSPMTSLYK	739
CC	The present invention relates to the isolation of novel human enzyme	RESULT 9			
CC	polypeptides, and the cDNA (AAS40785-AAS41684) and genomic sequences	AA919154			
CC	encoding them. The enzyme polypeptides of the invention may comprise the	ID	AA919154	standard; protein; 731 AA.	
CC	functional classes of oxidoreductases, transferases, hydrolases, lyases,	XX	AA919154		
CC	isomerases or ligases. The sequences of the invention are useful in the	AC	AA919154		
CC	diagnosis, treatment, prevention and/or prognosis of a wide range of	XX			
CC	disorders including hyperproliferative disorders (e.g. cancer),	DT	19-JUL-2000	(first entry) - late	
CC	immunodeficiency disorders (e.g. AIDS) autoimmune disorders (e.g.	XX			
CC	arthritis), neurological disorders (e.g. Alzheimer's disease), metabolic	DE		Human cytoskeleton associated protein 9 (CYSKP-9).	
CC	disorders (e.g. phenylketonuria), inflammatory disorders (e.g. asthma),	XX			
CC	cardiovascular disorders (e.g. atherosclerosis), blood-related disorders	KW		Cytoskeleton associated protein; CYSKP-9; cancer; proliferative;	
CC	(e.g. haemophilia), reproductive disorders (e.g. infertility) and	KW		autoimmunity; inflammatory, vesicle trafficking; neurological;	
CC	infectious disorders (e.g. influenza). The polynucleotides of the	KW		cardiovascular; cell motility; reproductive; muscle disorder.	
CC	invention can also be used in gene therapy. AAU22915-AAU23814 represent	XX			
CC	the novel human enzyme polypeptides of the invention. Note: The sequence	OS		Homo sapiens.	
CC	data for this patent did not form part of the printed specification, but	XX			
CC	was obtained in electronic format directly from WIPO at	FT	Key	Location/Qualifiers	
CC	ftp.wipo.int/pub/published_pct_sequences	FT	Modified-site	15	
XX	Sequence 739 AA;	FT	Domain	40..64	/note= "potential phosphorylation site"
SQ		FT	Modified-site	117	/note= "CAP-Gly domain"
		FT	Modified-site	125	/note= "potential phosphorylation site"
		FT	Modified-site	131	/note= "potential glycosylation site"
		FT	Modified-site	134	/note= "potential phosphorylation site"
		FT	Modified-site	136	/note= "potential glycosylation site"
		FT	Modified-site	162	/note= "potential phosphorylation site"
		FT	Modified-site	168	/note= "potential phosphorylation site"
		FT	Modified-site	205	/note= "potential phosphorylation site"
		FT	Modified-site	219	/note= "potential glycosylation site"
		FT	Modified-site	249	/note= "potential phosphorylation site"
		FT	Modified-site	292	/note= "potential phosphorylation site"
		FT	Modified-site	313	/note= "potential phosphorylation site"
		FT	Modified-site	390	/note= "potential phosphorylation site"
		FT	Modified-site	407	/note= "potential phosphorylation site"
		FT	Modified-site	451	/note= "potential phosphorylation site"
		FT	Modified-site	490	/note= "potential phosphorylation site"

PR	11-DEC-2000; 2000US-0254097P.	QY	661	KVEAASGFTSEKDPPEFLNLPFHLLRVPEPLKIRSAGQVQDCYFYQIFMEKNEKVG	720
PR	05-JAN-2001; 2001US-0259678P.	Db	451	KVEAASGFTSEKDPPEFLNLPFHLLRVPEPLKIRSAGQVQDCYFYQIFMEKNEKVG	510
PA	(HUMA-) HUMAN GENOME SCI INC.	QY	721	PTIQQLLEWSFINSNLKFAEAPSCLIIQMPFRGKDFKLFKKI PPSLELNTDLEDTPRQ	780
XX	Rosen CA, Barash SC, Ruben SM;	Db	511	PTIQQLLEWSFINSNLKFAEAPSCLIIQMPFRGKDFKLFKKI PPSLELNTDLEDTPRQ	570
XX	WPI, 2001-465566/50.	QY	781	CRICGGLAMVECRECYDDPDISAGIKQFCKTQNTQVHLHPKRLNHKYNPVSPLPKDLPDW	840
DR	N-PSDB; AAS41617.	Db	571	CRICGGLAMVECRECYDDPDISAGIKQFCKTQNTQVHLHPKRLNHKYNPVSPLPKDLPDW	630
PT	Novel polypeptides and polynucleotides useful for diagnosing, preventing,	QY	841	DWRHGCIPQCNMELFAVLCTIETSHYVAFVKYKDDSAWLPFDSMADRDGGONGNIPOVT	900
PT	treating neural, immune system, muscular, reproductive, pulmonary,	Db	631	DWRHGCIPQCNMELFAVLCTIETSHYVAFVKYKDDSAWLPFDSMADRDGGONGNIPOVT	690
PT	cardiovascular, renal, proliferative disorders and cancerous diseases.	QY	901	PCPEVGEYLKMSLEDLSRRIOGCARRLLCDAYCMYQSPMTSLYK	949
XX	Claim 11; SEQ ID NO 1743; 1180pp; English.	Db	691	PCPEVGEYLKMSLEDLSRRIOGCARRLLCDAYCMYQSPMTSLYK	739
CC	The present invention relates to the isolation of novel human enzyme	RESULT 9			
CC	polypeptides, and the cDNA (AAS40785-AAS41684) and genomic sequences	AA919154			
CC	encoding them. The enzyme polypeptides of the invention may comprise the	ID	AA919154	standard; protein; 731 AA.	
CC	functional classes of oxidoreductases, transferases, hydrolases, lyases,	XX	AA919154		
CC	isomerases or ligases. The sequences of the invention are useful in the	AC	AA919154		
CC	diagnosis, treatment, prevention and/or prognosis of a wide range of	XX			
CC	disorders including hyperproliferative disorders (e.g. cancer),	DT	19-JUL-2000	(first entry) - late	
CC	immunodeficiency disorders (e.g. AIDS) autoimmune disorders (e.g.	XX			
CC	arthritis), neurological disorders (e.g. Alzheimer's disease), metabolic	DE		Human cytoskeleton associated protein 9 (CYSKP-9).	
CC	disorders (e.g. phenylketonuria), inflammatory disorders (e.g. asthma),	XX			
CC	cardiovascular disorders (e.g. atherosclerosis), blood-related disorders	KW		Cytoskeleton associated protein; CYSKP-9; cancer; proliferative;	
CC	(e.g. haemophilia), reproductive disorders (e.g. infertility) and	KW		autoimmunity; inflammatory, vesicle trafficking; neurological;	
CC	infectious disorders (e.g. influenza). The polynucleotides of the	KW		cardiovascular; cell motility; reproductive; muscle disorder.	
CC	invention can also be used in gene therapy. AAU22915-AAU23814 represent	XX			
CC	the novel human enzyme polypeptides of the invention. Note: The sequence	OS		Homo sapiens.	
CC	data for this patent did not form part of the printed specification, but	XX			
CC	was obtained in electronic format directly from WIPO at	FT	Key	Location/Qualifiers	
CC	ftp.wipo.int/pub/published_pct_sequences	FT	Modified-site	15	
XX	Sequence 739 AA;	FT	Domain	40..64	/note= "potential phosphorylation site"
SQ		FT	Modified-site	117	/note= "CAP-Gly domain"
		FT	Modified-site	125	/note= "potential phosphorylation site"
		FT	Modified-site	131	/note= "potential glycosylation site"
		FT	Modified-site	134	/note= "potential phosphorylation site"
		FT	Modified-site	136	/note= "potential glycosylation site"
		FT	Modified-site	162	/note= "potential phosphorylation site"
		FT	Modified-site	168	/note= "potential phosphorylation site"
		FT	Modified-site	205	/note= "potential phosphorylation site"
		FT	Modified-site	219	/note= "potential glycosylation site"
		FT	Modified-site	249	/note= "potential phosphorylation site"
		FT	Modified-site	292	/note= "potential phosphorylation site"
		FT	Modified-site	313	/note= "potential phosphorylation site"
		FT	Modified-site	390	/note= "potential phosphorylation site"
		FT	Modified-site	407	/note= "potential phosphorylation site"
		FT	Modified-site	451	/note= "potential phosphorylation site"
		FT	Modified-site	490	/note= "potential phosphorylation site"

QY	184	QLFQCDDECG-FVALDKLELIEDDDESDYAGPDTMQLPPLNRSVLSKGETI	242
Db	1	QLFQCDDECGFVALDKLELIEDDDESDYAGPDTMQLPPLNRSVLSK-----	55
QY	243	ESGTVIFCDVLPKESGLGVGVGDMNDPIGNWDRFDGV-LCSFACVESTILLHNDIIP	301
Db	56	-----DNPIGNWDRFDGVQLCSFACVESTILLHNDIIP	90
QY	302	ESVTOERRPPKLAFLMSRGVKGSSSHNKPATGTSDFGNR-RSELFYTLNGSSVDSQP	360
Db	91	ESVTOERRPPKLAFLMSRGVKGSSSHNKPATGTSDFGNRSELFYTLNGSSVDSQP	150
QY	361	QSKSKNTWYIDEVADPAKSLTEISTDFDRSPPLQPPVNSLTENRPHSLPFLTKMP	420
Db	151	QSKSKNTWYIDEVADPAKSLTEISTDFDRSPPLQPPVNSLTENRPHSLPFLTKMP	210
QY	421	NTNGSIGHSPLSLSAQSVMEELNTPVQESPLAMPNGSHGLEVGS LAEVKENPPFYGV	480
Db	211	NTNGSIGHSPLSLSAQSVMEELNTPVQESPLAMPNGSHGLEVGS LAEVKENPPFYGV	270
QY	481	IRWIGQPPGLNEVLAGLEDEDCAGTDCGTFRGTRYFTCALKKALFVKLSCRPDRFAS	540
Db	271	IRWIGQPPGLNEVLAGLEDEDCAGTDCGTFRGTRYFTCALKKALFVKLSCRPDRFAS	330
QY	541	LQPVSNQIERCNLSLAFGGVLSVWBEENTPPKMEKGLIEMI GKKGIOGHYNSCYLDSTL	600
Db	331	LQPVSNQIERCNLSLAFGGVLSVWBEENTPPKMEKGLIEMI GKKGIOGHYNSCYLDSTL	390
QY	601	FCLFAPSSVLDTVLLRPKEKNDVEYVETQELLRTIENVPLRIYGVVCA TKIMLRKILE	660
Db	391	FCLFAPSSVLDTVLLRPKEKNDVEYVETQELLRTIENVPLRIYGVVCA TKIMLRKILE	450

Query Match 77.0%; Score 3876.5; DB 4; Length 739;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 736; Conservative 0; Mismatches 0; Indels 33; Gaps 4;

FT	Modified-site	/note= "potential phosphorylation site"	Db	241	NSHGLEVGSLAEVKENPPYGVIRWIGPPGLNEVLAGLEDEACAGCTDGTGRTRYFT	300
FT	Modified-site	/note= "potential glycosylation site"	Qy	519	CALKKALFVKLXCRPDSRFASLQPVSNQIERCNSLAFGGYLSVVVEENTPPKMEKEGLE	578
FT	Modified-site	/note= "potential phosphorylation site"	Db	301	CALKKALFVKLXCRPDSRFASLQPVSNQIERCNSLAFGGYLSVVVEENTPPKMEKEGLE	360
FT	Modified-site	/note= "potential phosphorylation site"	Qy	579	IMIGKKGIQGHYNSCYLDSTLFLCLFAPSSVLDTVLLRPKEKNDVEYSETQELLRTIIV	638
FT	Modified-site	/note= "potential phosphorylation site"	Db	361	IMIGKKGIQGHYNSCYLDSTLFLCLFAPSSVLDTVLLRPKEKNDVEYSETQELLRTIIV	420
FT	Modified-site	/note= "potential phosphorylation site"	Qy	639	NPLRIYGVYCATKIMKRLKILEKVEAASGFTSEKDPPEEFNLIFHILRVEPLLKIRSA	698
FT	Modified-site	/note= "potential phosphorylation site"	Db	421	NPLRIYGVYCATKIMKRLKILEKVEAASGFTSEKDPPEEFNLIFHILRVEPLLKIRSA	480
Qy			Qy	699	GQKVQDCYFYQIFMEKNEKVGVTPTIQQLLEWFSFINSNLKFAEAPSCLIIQMPPFGKDFKL	758
Db			Db	481	GQKVQDCYFYQIFMEKNEKVGVTPTIQQLLEWFSFINSNLKFAEAPSCLIIQMPPFGKDFKL	540
Qy			Qy	759	FKKIPPSLELNTDLLEDTTPROCRICGGLAMVECRECYDDPDISAGIKIQFCKTCNTQVH	818
Db			Db	541	FKKIPPSLELNTDLLEDTTPROCRICGGLAMVECRECYDDPDISAGIKIQFCKTCNTQVH	600
Qy			Qy	819	LHPKRLNHKYNPVSPLKDLPDMDWRHGCI PCQNMELFAVL CIETSHVAFVKYKDDSAW	878
Db			Db	601	LHPKRLNHKYNPVSPLKDLPDMDWRHGCI PCQNMELFAVL CIETSHVAFVKYKDDSAW	660
Qy			Qy	879	LFDFSMDRGGQNGFNIPQVTPCPVEGYLXMSLEDHLSLDSRRRIQGCARRLLCDAYMC	938
Db			Db	661	LFDFSMDRGGQNGFNIPQVTPCPVEGYLXMSLEDHLSLDSRRRIQGCARRLLCDAYMC	720
Qy			Qy	939	MYQSPMTSLYK 949	
Db			Db	721	MYQSPMTSLYK 731	
RESULT 10						
AAB95828						
ID	AAB95828 standard; protein; 731 AA.					
XX						
AC	AAB95828;					
DT	26-JUN-2001 (first entry) - late					
DE	Human protein sequence SEQ ID NO:18843.					
XX						
KW	Human; primer; detection; diagnosis; antisense therapy; gene therapy.					
OS	Homo sapiens.					
XX						
PN	EP1074617-A2.					
XX						
PD	07-FEB-2001.					
XX						
PF	28-JUL-2000; 2000EP-00116126.					
XX						
PR	29-JUL-1999; 99JP-00248036.					
PR	27-AUG-1999; 99JP-00300253.					
PR	11-JAN-2000; 2000JP-00118776.					
PR	02-MAY-2000; 2000JP-00183767.					
PR	09-JUN-2000; 2000JP-00241899.					
XX	(HELI-) HELIX RES INST.					
XX						
PI	Ota T, Isogai T, Nishikawa T, Hayaishi K, Saito K, Yamamoto J;					
PI	Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;					
XX	WPI; 2001-318749/34.					
XX						
PT	Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.					
PT						
XX						

no 102(e) date allowed

Human cytoskeleton associated proteins, used to treat cell proliferative, autoimmune/inflammatory, vesicle trafficking, neurological, cell motility, reproductive and muscle disorders.

Claim 1; Page 91-93; 113pp; English.

AA91946-61 show human cytoskeleton associated proteins 1 to 16 (CYSKP-1 to CYSKP-16) respectively. The sequences can be used to treat and diagnose cancer and cell proliferative, autoimmune/inflammatory, vesicle trafficking, neurological, cardiovascular, cell motility, reproductive and muscle disorders. Pharmaceutical compositions containing CYSKP-2 to CYSKP-16 can be used to treat or prevent disorders associated with decreased expression or activity of CYSKP (claimed), for example, atherosclerosis, cirrhosis, hepatitis, myelofibrosis, psoriasis, cancers, autoimmune/antimflammatory disorders such as allergies, anemia, asthma, acquired immunodeficiency syndrome (AIDS), Crohn's disease, diabetes mellitus, Goodpasture's syndrome, osteoarthritis, scleroderma and trauma. CYSKP antagonists can be used to treat or prevent a disorder associated with increased expression or activity of CYSKP (claimed)

Query Match 76.7%; Score 3862; DB 3; Length 731; Best Local Similarity 99.6%; Pred. No. 0; Matches 728; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

Qy 221 MQVELPPEINSRVSLKGGTIESGVIFCDVLPKGESLGYFVGVDMDNPIGNWDRFGD 280

Db 1 MQVELPPEINSRVSLKGGTIESGVIFCDVLPKGESLGYFVGVDMDNPIGNWDRFGD 60

Qy 281 V-LCSFACVESTILLHINDIIPESVTQERRPPKLAFLAFMSRGVGDGKSSSHNKPKATGSTSD 339

Db 61 VOLCSFACVESTILLHINDIIPESVTQERRPPKLAFLAFMSRGVGDGKSSSHNKPKATGSTSD 120

Qy 340 PGNR-RSELYFTYNGSSVDSQPSKSNWYIDEVAEDPAKSLTEISTDFDRSSPPLOPP 398

Db 121 PGNRRSELYFTYNGSSVDSQPSKSNWYIDEVAEDPAKSLTEISTDFDRSSPPLOPP 180

Qy 399 PVSNTTENNRFSLPPLTKMPTNGSIGHSPLSLSAQSWMEELNAPVQESPPLAMPQG 458

Db 181 PVSNTTENNRFSLPPLTKMPTNGSIGHSPLSLSAQSWMEELNAPVQESPPLAMPQG 240

Qy 459 NSHGLEVGSLAEVKENPPYGVIRWIGPPGLNEVLAGLEDEACAGCTDGTGRTRYFT 518

PS	Claim 8; SEQ ID NO 18843; 2537pp + Sequence Listing; English.	
XX		
CC	The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises:	
CC	(a) an oligo-dr primer and an oligonucleotide complementary to the	
CC	complementary strand of a polynucleotide which comprises one of the 5602	
CC	nucleotide sequences defined in the specification, where the	
CC	oligonucleotide comprises at least 15 nucleotides; or (b) a combination	
CC	of an oligonucleotide comprising a sequence complementary to the	
CC	complementary strand of a polynucleotide which comprises a 5'-end	
CC	sequence and an oligonucleotide comprising a sequence complementary to a	
CC	polynucleotide which comprises a 3'-end sequence, where the	
CC	oligonucleotide comprises at least 15 nucleotides and the combination of	
CC	the 5'-end sequence/3'-end sequence is selected from those defined in the	
CC	specification. The primer sets can be used in antisense therapy and in	
CC	gene therapy. The primers are useful for synthesizing polynucleotides,	
CC	particularly full-length cDNAs. The primers are also useful for the	
CC	detection and/or diagnosis of the abnormality of the proteins encoded by	
CC	the full-length cDNAs. The primers allow obtaining of the full-length	
CC	cDNAs easily without any specialised methods. AAH03166 to AAH13628 and	
CC	AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893	
CC	represent human amino acid sequences; and AAH13629 to AAH13632 represent	
CC	oligonucleotides, all of which are used in the exemplification of the	
XX	present invention	
XX		
SQ	Sequence 731 AA;	
	Query Match 76.6%; Score 3855; DB 4; Length 731;	
	Best Local Similarity 99.5%; Pred. No. 0;	
	Matches 727; Conservative 0; Mismatches 2; Indels 2; Gaps 2;	
QY	221 MQVELPPLLEINSRVSLKGGTETIESGTVIFCDVLPGKESLGYFVGVDMDNPNIGNWGRFDG 280	
DB	1 MQVELPPLLEINSRVSLKGGTETIESGTVIFCDVLPGKESLGYFVGVDMDNPNIGNWGRFDG 60	
QY	281 V-LCSFACVESTILLHINDIIPESVTOERRPPKLAFMRSRGVGDGSSHNKPKATGSTSD 339	
DB	61 VQLCSFACVESTILLHINDIIPESVTOERRPPKLAFMRSRGVGDGSSHNKPKATGSTSD 120	
QY	340 PGNR-RSELFYTLNGSSVDSQPSKSNWTYIDEVAEDPAKSLTEISTDFDRSSPPLQPP 398	
DB	121 PGNR-RSELFYTLNGSSVDSQPSKSNWTYIDEVAEDPAKSLTEISTDFDRSSPPLQPP 180	
QY	399 PVNSLTENRPHSLPPLSLTKMPNTNGSIGHSPLSLQAQSVMEELNAPVQESPLAMPFG 458	
DB	181 PVNSLTENRPHSLPPLSLTKMPNTNGSIGHSPLSLQAQSVMEELNAPVQESPLAMPFG 240	
QY	459 NSHGLEVGSLAEVKENPPYGVIRWTGPPGLNEVLVAGLEDEACAGCTDGTFRGTRYFT 518	
DB	241 NSHGLEVGSLAEVKENPPYGVIRWTGPPGLNEVLVAGLEDEACAGCTDGTFRGTRYFT 300	
QY	519 CALKKALFVKLAKSRPDSRFASIQPVSNQIERCNLSIAFGYLSVVEENTPPKMEKEGLE 578	
DB	301 CALKKALFVKLAKSRPDSRFASIQPVSNQIERCNLSIAFGYLSVVEENTPPKMEKEGLE 360	
QY	579 IMIGKKGIQGHYNSCYLSDTLFCLFAPFSVLDTVLLRPKEKNDVYVYSETQBLRLTEIV 638	
DB	361 IMIGKKGIQGHYNSCYLSDTLFCLFAPFSVLDTVLLRPKEKNDVYVYSETQBLRLTEIV 420	
QY	639 NPLRIYGYCATKIMKRLKILEKVEAASGFTSEKDPPEBFLNLFPHILRVEPLLKIRSA 698	
DB	421 NPLRIYGYCATKIMKRLKILEKVEAASGFTSEKDPPEBFLNLFPHILRVEPLLKIRSA 480	
QY	699 GQKVQDCYFYQIFMEKNEKVGPTIQOLLEWSFINSNLKFAEAPSLIOMPRFGKDFKL 758	
DB	481 GQKVQDCYFYQIFMEKNEKVGPTIQOLLEWSFINSNLKFAEAPSLIOMPRFGKDFKL 540	
QY	759 FKKIPPSLENTDLDLTPROCRICGGGLAMVRECYDDPDISAGKIKQFCKTCTQVH 818	
DB	541 FKKIPPSLENTDLDLTPROCRICGGGLAMVRECYDDPDISAGKIKQFCKTCTQVH 600	
QY	819 LHPKRLNHKNYPVSLPKDLPDMDWRHGCCIPQNMELFAVLICITSHYVAFVKYKDDSAW 878	
DB	601 LHPKRLNHKNYPVSLPKDLPDMDWRHGCCIPQNMELFAVLICITSHYVAFVKYKDDSAW 660	
QY	879 LFPDSDMADRDGGQNGFNIPQVTPCPEVGEYLRKMSLEDLHSLDSRRIOGCARRLLCDAYMC 938	
DB	661 LFPDSDMADRDGGQNGFNIPQVTPCPEVGEYLRKMSLEDLHSLDSRRIOGCARRLLCDAYMC 720	
QY	939 MYQSPTMSLYK 949	
DB	721 MYQSPTMSLYK 731	
	RESULT 11	
AM39254		
ID	AAH39254 standard; protein; 685 AA.	
XX		
AC	AAH39254;	
XX		
DT	22-OCT-2001 (first entry) - <i>late</i>	
XX		
DE	Human polypeptide SEQ ID NO 2399.	
XX		
KW	Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;	
KW	peripheral nervous system; neuropathy; central nervous system; CNS;	
KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;	
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;	
KW	chemokinetic; thrombolytic; drug screening; arthritis; inflammation;	
KW	leukaemia.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200153312-A1.	
XX		
PD	26-JUL-2001.	
XX		
PF	26-DEC-2000; 2000WO-US034263.	
XX		
PR	23-DEC-1999; 99US-00471275.	
PR	21-JAN-2000; 2000US-00488725.	
PR	25-APR-2000; 2000US-0052317.	
PR	20-JUN-2000; 2000US-00598042.	
PR	19-JUL-2000; 2000US-00620312.	
PR	03-AUG-2000; 2000US-00653450.	
PR	14-SEP-2000; 2000US-00662191.	
PR	19-OCT-2000; 2000US-00693036.	
PR	29-NOV-2000; 2000US-00727344.	
XX	(HYSE-) HYSEQ INC.	
XX		
PI	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;	
PI	Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;	
PI	Zhou P, Goodrich R, Drmanac RT;	
XX		
DR	WPI; 2001-442253/47.	
DR	N-PSDB; AA158410.	
XX		
PT	Novel nucleic acids and polypeptides, useful for treating disorders such	
PT	as central nervous system injuries.	
XX		
PS	Example 4; SEQ ID NO 2399; 10078pp; English.	
XX		
CC	The invention relates to human nucleic acids (AAI57798-AAI61369) and the	
CC	encoded polypeptides (AAH38642-AAH42213) with nootropic,	
CC	immunosuppressant and cytostatic activity. The polynucleotides are useful	
CC	in gene therapy. A composition containing a polypeptide or polynucleotide	
CC	of the invention may be used to treat diseases of the peripheral nervous	
CC	system, such as peripheral nervous injuries, peripheral neuropathy and	
CC	localised neuropathies and central nervous system diseases, such as	
CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic	
CC	lateral sclerosis, and Shy-Drager Syndrome. Other uses include the	
CC	utilisation of the activities such as: Immune system suppression,	
CC	Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic	
CC	and thrombolytic activity, cancer diagnosis and therapy, drug screening,	
CC	assays for receptor activity, arthritis and inflammation, leukaemias and	

CC	C.N.S disorders. Note: The sequence data for this patent did not form part of the printed specification	
XX		
SQ	Sequence 685 AA;	
	Query Match 72.2%; Score 3636; DB 4; Length 685;	
	Best Local Similarity 99.7%; Pred. No. 0;	
	Matches 683; Conservative 0; Mismatches 0; Indels 2; Gaps 2;	
Qy	267 MDNPIGNWDRFDGV-LCSFACVESTILLHINDIIPESVTQERRPPKLAFLMRGVDGKGS 325	
Db	1 MDNPIGNWDRFDGV-LCSFACVESTILLHINDIIPESVTQERRPPKLAFLMRGVDGKGS 60	
Qy	326 SSHNPKATGSTDGPNR-RSELFYTLNGSSVDSPQSKNKTWYIDEVAEDPAKSLTEI 384	
Db	61 SSHNPKATGSTDGPNRSELFYTLNGSSVDSPQSKNKTWYIDEVAEDPAKSLTEI 120	
Qy	395 STDPRSSPPLOPPVNSLTNNRPHSLPFLSTKMPNTNGSIGHSPLSLSAQSVMEELNT 444	
Db	121 STDPRSSPPLOPPVNSLTNNRPHSLPFLSTKMPNTNGSIGHSPLSLSAQSVMEELNT 180	
Qy	445 APVQSPPLAMPNGSHGLEVSLAEVKENPPFYGVIRWIGOPGLNEVLAGELEDECA 504	
Db	181 APVQSPPLAMPNGSHGLEVSLAEVKENPPFYGVIRWIGOPGLNEVLAGELEDECA 240	
Qy	505 GCTDGTFRGTRYFTCALKKALFVKLSCRPDSRFASLQPVSNQIERCNSLAFGGYLSVV 564	
Db	241 GCTDGTFRGTRYFTCALKKALFVKLSCRPDSRFASLQPVSNQIERCNSLAFGGYLSVV 300	
Qy	565 BENTPPKMEKEGLEIMIGKKGIQGHYNSCYLDSTLFCFLFAPSSVLDVLLRPKEKNVE 624	
Db	301 BENTPPKMEKEGLEIMIGKKGIQGHYNSCYLDSTLFCFLFAPSSVLDVLLRPKEKNVE 360	
Qy	625 YVSETQELLRTIEVNPRLTYGVVCAKIMKLKILEKVEAASGFTSEKDPDEFNLIRPH 684	
Db	361 YVSETQELLRTIEVNPRLTYGVVCAKIMKLKILEKVEAASGFTSEKDPDEFNLIRPH 420	
Qy	685 HILRVEPLLKIRSAGQKQVDCYFYQIFMEKNEKVGVPYTIQQLLEWSFINSNLKFAEAPSC 744	
Db	421 HILRVEPLLKIRSAGQKQVDCYFYQIFMEKNEKVGVPYTIQQLLEWSFINSNLKFAEAPSC 480	
Qy	745 LIIQWPRFGOKPLFKKIPPSLELNITDLETPRQCRIICGLAMYECREYDDPDISAG 804	
Db	481 LIIQWPRFGOKPLFKKIPPSLELNITDLETPRQCRIICGLAMYECREYDDPDISAG 540	
Qy	805 KIKQFCKTNTQVHLHPKELNHYNPVSLPKDLPDMDWRHGCTPCQNMELFAVLCTETSH 864	
Db	541 KIKQFCKTNTQVHLHPKELNHYNPVSLPKDLPDMDWRHGCTPCQNMELFAVLCTETSH 600	
Qy	865 YVAFVKYKDDSAWLFFDSDMADRGGQGNFNPQVTPCPVEGYLKMSLDLHSLDSRI 924	
Db	601 YVAFVKYKDDSAWLFFDSDMADRGGQGNFNPQVTPCPVEGYLKMSLDLHSLDSRI 560	
Qy	925 QGCARLLCDAYMCMYQSTMSLYK 949	
Db	661 QGCARLLCDAYMCMYQSTMSLYK 685	
RESULT 12		
AA41040		
ID	AA41040 standard; protein; 698 AA.	
XX		
AC	AA41040;	
XX		
XX	22-OCT-2001 (first entry)	
DT	Human polypeptide SEQ ID NO 5971.	
XX		
DE		
XX		
KW	Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;	
KW	peripheral nervous system; neuropathy; central nervous system; CNS;	
KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;	
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;	
KW	chemokinetic; thrombolytic; drug screening; arthritis; inflammation;	

KW	leukaemia.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200153312-A1.	
XX		
PD	26-JUL-2001.	
XX		
PF	26-DEC-2000; 2000WO-US034263.	
XX		
PR	23-DEC-1999; 99US-00471275.	
PR	21-JAN-2000; 2000US-00488725.	
PR	25-APR-2000; 2000US-0052317.	
PR	20-JUN-2000; 2000US-00598042.	
PR	19-JUL-2000; 2000US-00620312.	
PR	03-AUG-2000; 2000US-00653450.	
PR	14-SEP-2000; 2000US-00662191.	
PR	19-OCT-2000; 2000US-00693036.	
PR	29-NOV-2000; 2000US-00727344.	
XX		
PA	(HYSE-) HYSEQ INC.	
XX		
PI	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;	
PI	Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;	
PI	Zhou P, Goodrich R, Drmanac RT;	
XX		
DR	WPI; 2001-442253/47.	
DR	N-PSDB; AA160196.	
XX		
XX	Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries.	
PS	Example 2; SEQ ID NO 5971; 10078pp; English.	
XX		
CC	The invention relates to human nucleic acids (AA157798-AA161369) and the encoded polypeptides (AA38642-AA42213) with nootropic,	
CC	immunosuppressant and cytostatic activity. The polynucleotides are useful in gene therapy. A composition containing a polypeptide or polynucleotide of the invention may be used to treat diseases of the peripheral nervous system, such as peripheral nervous injuries, peripheral neuropathy and localised neuropathies and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the utilisation of the activities such as: Immune system suppression,	
CC	Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, cancer diagnosis and therapy, drug screening, assays for receptor activity, arthritis and inflammation, leukaemias and C.N.S disorders. Note: The sequence data for this patent did not form part of the printed specification	
XX		
SQ	Sequence 698 AA;	
	Query Match 70.8%; Score 3563; DB 4; Length 698;	
	Best Local Similarity 97.8%; Pred. No. 5.2e-314;	
	Matches 680; Conservative 2; Mismatches 7; Indels 6; Gaps 6;	
Qy	261 YFVGVDMDNPIGNWDRFDGV-LCSFACVESTILLHINDIIPESVTQERRPPKLAFLMRG 319	
Db	4 YFVGVDMDNPIGNWDRFDGV-LCSFACVESTILLHINDIIPESVTQERRPPKLAFLMRG 63	
Qy	320 VDGKSSSHNPKATGSTDGPNR-RSELFYTLNGSSVDSPQSKNKTWYIDEVAEDPA 378	
Db	64 VDGKSSSHNPKATGSTDGPNRSELFYTLNGSSVDSPQSKNKTWYIDEVAEDPA 123	
Qy	379 KSLTEISTDFDRSSPPLQPPVNSLTNNRPHSLPFLSTKMPNTNGSIGHSPLSAQSV 438	
Db	124 KSLTEISTDFDRSSPPLQPPVNSLTNNRPHSLPFLSTKMPNTNGSIGHSPLSAQSV 183	
Qy	439 MEELNTAPVQSPPLAMPNGSHGLEVSLAEVKENPPFYGVIRWIGOPGLNEVLAGE 498	
Db	184 MEELNTAPVQSPPLAMPNGSHGLEVSLAEVKENPPFYGVIRWIGOPGLNEVLAGE 243	
Qy	499 LEDECAGCTDGTFR-GTRYFTCALKKALFVKLSCRPDSRFASLQPVSNQIERCNSLA-F 556	

Db	244	LEDEAGCTDGTREGTRYFTCALKKALFVKLKSRCPSDFASIQPVNSQIERCNSLAIW	303	PR	26-JUL-2000;	2000US-0220963P.
Qy	557	GGYLSEVVENPT-PRMEKEGLEIMIGKKGIQGHYNSCYLDSTLFCLEAFSSVLDTVLL	615	PR	26-JUL-2000;	2000US-0220964P.
Db	304	EAYLSEVVEENTPTQWEKEGLEIMIGKKGIQGHYNSCYLDSTLFCLEAFSSVLDTVLL	363	PR	14-AUG-2000;	2000US-0224518P.
Qy	616	RPKEKNDVEYSTQELLRTIENVPLRIYGVVCAKIMKRLKILEKVEAASGFTSEBKDP	675	PR	14-AUG-2000;	2000US-0225213P.
Db	364	RPKEKNDVEYSTQELLRTIENVPLRIYGVVCAKIMKRLKILEKVEAASGFTSEBKDP	423	PR	14-AUG-2000;	2000US-0225214P.
Qy	676	EEFLNLPHHILRVEPLLKIRSAQKQVQDCYFYQIFMEKNEKVGVPITIQOLLEWSFNSN	735	PR	14-AUG-2000;	2000US-0225266P.
Db	424	EEFLNLPHHILRVEPLLKIRSAQKQVQDCYFYQIFMEKNEKVGVPITIQOLLEWSFNSN	483	PR	14-AUG-2000;	2000US-0225267P.
Qy	736	LKFAEAPSLIIOMPRFGKDFKFKIPPSLELNITDLEDTPRQCRICGGLAMEYCREC	795	PR	14-AUG-2000;	2000US-0225268P.
Db	484	LKFAEAPSLIIOMPRFGKDFKFKIPPSLELNITDLEDTPRQCRICGGLAMEYCREC	543	PR	14-AUG-2000;	2000US-0225270P.
Qy	796	YDDPDISAGKIKOFCCTQNTQVHLHPKRLNHNKPNVSLPKDLPDWDWRHGCIPQNNMELF	855	PR	14-AUG-2000;	2000US-0225757P.
Db	544	YDDPDISAGKIKOFCCTQNTQVHLHPKRLNHNKPNVSLPKDLPDWDWRHGCIPQNNMELF	603	PR	14-AUG-2000;	2000US-0225758P.
Qy	856	AVLCIETSHYVAFVKYCKDSSAWLFFDSMADRDGGQNGFNIPQVTPCPEVGEYLKMSLED	915	PR	14-AUG-2000;	2000US-0225759P.
Db	604	AVLCIETSHYVAFVKYCKDSSAWLFFDSMADRDGGQNGFNIPQVTPCPEVGEYLKMSLED	663	PR	18-AUG-2000;	2000US-0226279P.
Qy	916	LHSLDERRIQGCARRLLCDA-YNCMTQSPMTSLYK	949	PR	22-AUG-2000;	2000US-0226681P.
Db	664	LHSLDERRIQGCARRLLCDAIYVVPCTQSPMTSLYK	698	PR	22-AUG-2000;	2000US-0226681P.
RESULT 13					22-AUG-2000;	2000US-0226686P.
AAU23211					23-AUG-2000;	2000US-0227182P.
ID	AAU23211	standard; protein; 558 AA.			23-AUG-2000;	2000US-0227009P.
XX	AAU23211;				30-AUG-2000;	2000US-0228942P.
XX	AAU23211;				01-SEP-2000;	2000US-0229343P.
DT	18-DEC-2001	(first entry)			01-SEP-2000;	2000US-0229344P.
DE	Novel human enzyme polypeptide #297.				01-SEP-2000;	2000US-0229345P.
XX	Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;				05-SEP-2000;	2000US-0229509P.
KW	ligase; hyperproliferative disorder; immunodeficiency disorder;				05-SEP-2000;	2000US-0229509P.
KW	autoimmune disorder; neurological disorder; metabolic disorder;				06-SEP-2000;	2000US-0230437P.
KW	inflammatory disorder; cardiovascular disorder; reproductive disorder;				06-SEP-2000;	2000US-0230438P.
KW	blood-related disorder; infectious disorder; cytostatic; anti arthritic;				08-SEP-2000;	2000US-0231242P.
XX	nephrotropic; anticoagulant.				08-SEP-2000;	2000US-0231243P.
OS	Homo sapiens.				08-SEP-2000;	2000US-0231244P.
XX	WO200155301-A2.				08-SEP-2000;	2000US-0231413P.
PN	WO200155301-A2.				08-SEP-2000;	2000US-0231414P.
XX	02-AUG-2001.				08-SEP-2000;	2000US-0232080P.
XX	17-JAN-2001; 2001WO-US001239.				08-SEP-2000;	2000US-0232081P.
XX	31-JAN-2000; 2000US-0179065P.				12-SEP-2000;	2000US-0231968P.
PR	04-FEB-2000; 2000US-0180628P.				14-SEP-2000;	2000US-0232398P.
PR	24-FEB-2000; 2000US-0184664P.				14-SEP-2000;	2000US-0232399P.
PR	02-MAR-2000; 2000US-0186350P.				14-SEP-2000;	2000US-0232400P.
PR	16-MAR-2000; 2000US-0189874P.				14-SEP-2000;	2000US-0232401P.
PR	17-MAR-2000; 2000US-0190076P.				14-SEP-2000;	2000US-0233063P.
PR	18-APR-2000; 2000US-0198123P.				14-SEP-2000;	2000US-0233064P.
PR	19-MAY-2000; 2000US-0205515P.				14-SEP-2000;	2000US-0233065P.
PR	07-JUN-2000; 2000US-0209467P.				21-SEP-2000;	2000US-0234223P.
PR	28-JUN-2000; 2000US-0214886P.				21-SEP-2000;	2000US-0234274P.
PR	30-JUN-2000; 2000US-0215135P.				25-SEP-2000;	2000US-0234997P.
PR	07-JUL-2000; 2000US-0216647P.				25-SEP-2000;	2000US-0234998P.
PR	07-JUL-2000; 2000US-0216880P.				26-SEP-2000;	2000US-0235484P.
PR	11-JUL-2000; 2000US-0217487P.				27-SEP-2000;	2000US-0235836P.
PR	14-JUL-2000; 2000US-0217496P.				29-SEP-2000;	2000US-0236327P.
PR	14-JUL-2000; 2000US-0218290P.				29-SEP-2000;	2000US-0236367P.
PR	14-JUL-2000; 2000US-0218290P.				29-SEP-2000;	2000US-0236368P.
PR	14-JUL-2000; 2000US-0218290P.				29-SEP-2000;	2000US-0236369P.
PR	14-JUL-2000; 2000US-0218290P.				02-OCT-2000;	2000US-0237037P.
PR	14-JUL-2000; 2000US-0218290P.				02-OCT-2000;	2000US-0237038P.
PR	14-JUL-2000; 2000US-0218290P.				02-OCT-2000;	2000US-0237039P.
PR	14-JUL-2000; 2000US-0218290P.				02-OCT-2000;	2000US-0237040P.
PR	14-JUL-2000; 2000US-0218290P.				13-OCT-2000;	2000US-0239935P.
PR	14-JUL-2000; 2000US-0218290P.				13-OCT-2000;	2000US-0239937P.
PR	14-JUL-2000; 2000US-0218290P.				20-OCT-2000;	2000US-0240960P.
PR	14-JUL-2000; 2000US-0218290P.				20-OCT-2000;	2000US-0241221P.
PR	14-JUL-2000; 2000US-0218290P.				20-OCT-2000;	2000US-0241785P.
PR	14-JUL-2000; 2000US-0218290P.				20-OCT-2000;	2000US-0241786P.
PR	14-JUL-2000; 2000US-0218290P.				20-OCT-2000;	2000US-0241787P.
PR	14-JUL-2000; 2000US-0218290P.				20-OCT-2000;	2000US-0241808P.
PR	14-JUL-2000; 2000US-0218290P.				20-OCT-2000;	2000US-0241809P.
PR	14-JUL-2000; 2000US-0218290P.				20-OCT-2000;	2000US-0241826P.
PR	14-JUL-2000; 2000US-0218290P.				01-NOV-2000;	2000US-0244617P.
PR	14-JUL-2000; 2000US-0218290P.				08-NOV-2000;	2000US-0244647P.

08-NOV-2000; 2000US-0246475P.
08-NOV-2000; 2000US-0246476P.
08-NOV-2000; 2000US-0246477P.
08-NOV-2000; 2000US-0246478P.
08-NOV-2000; 2000US-0246523P.
08-NOV-2000; 2000US-0246524P.
08-NOV-2000; 2000US-0246525P.
08-NOV-2000; 2000US-0246526P.
08-NOV-2000; 2000US-0246527P.
08-NOV-2000; 2000US-0246528P.
08-NOV-2000; 2000US-0246532P.
08-NOV-2000; 2000US-0246609P.
08-NOV-2000; 2000US-0246610P.
08-NOV-2000; 2000US-0246611P.
08-NOV-2000; 2000US-0246613P.
17-NOV-2000; 2000US-0249207P.
17-NOV-2000; 2000US-0249208P.
17-NOV-2000; 2000US-0249209P.
17-NOV-2000; 2000US-0249210P.
17-NOV-2000; 2000US-0249211P.
17-NOV-2000; 2000US-0249212P.
17-NOV-2000; 2000US-0249213P.
17-NOV-2000; 2000US-0249214P.
17-NOV-2000; 2000US-0249215P.
17-NOV-2000; 2000US-0249216P.
17-NOV-2000; 2000US-0249217P.
17-NOV-2000; 2000US-0249218P.
17-NOV-2000; 2000US-0249244P.
17-NOV-2000; 2000US-0249245P.
17-NOV-2000; 2000US-0249264P.
17-NOV-2000; 2000US-0249265P.
17-NOV-2000; 2000US-0249297P.
17-NOV-2000; 2000US-0249299P.
17-NOV-2000; 2000US-0249300P.
01-DEC-2000; 2000US-0250160P.
01-DEC-2000; 2000US-0250391P.
05-DEC-2000; 2000US-0251030P.
05-DEC-2000; 2000US-0251988P.
06-DEC-2000; 2000US-0251479P.
08-DEC-2000; 2000US-0251856P.
08-DEC-2000; 2000US-0251868P.
08-DEC-2000; 2000US-0251869P.
08-DEC-2000; 2000US-0251989P.
11-DEC-2000; 2000US-0254097P.
05-JAN-2001; 2001US-0259678P.
(HUNA-) HUMAN GENOME SCI INC.
Rosen CA, Barash SC, Ruben SM;
WPI; 2001-465566/50.
N-PSDB; AAS41081.
Novel polypeptides and polynucleotides useful for diagnosing, preventing, treating neural, immune system, muscular, reproductive, pulmonary, cardiovascular, renal, proliferative disorders and cancerous diseases.
Claim 11; SEQ ID NO 1207; 1180pp; English.
The present invention relates to the isolation of novel human enzyme polypeptides, and the cDNA (AAS40785-AAS41684) and genomic sequences encoding them. The enzyme polypeptides of the invention may comprise the functional classes of oxidoreductases, transferases, hydrolases, lyases, isomerases or ligases. The sequences of the invention are useful in the diagnosis, treatment, prevention and/or prognosis of a wide range of disorders, including hyperproliferative disorders (e.g. cancer), immunodeficiency disorders (e.g. AIDS) autoimmune disorders (e.g. arthritis), neurological disorders (e.g. Alzheimer's disease), metabolic disorders (e.g. phenylketonuria), inflammatory disorders (e.g. asthma), cardiovascular disorders (e.g. atherosclerosis), blood-related disorders (e.g. haemophilia), reproductive disorders (e.g. infertility) and

CC infectious disorders (e.g. Influenza). The polynucleotides of the invention can also be used in gene therapy. AAU22915-AAU23814 represent the novel human enzyme polypeptides of the invention. Note: The sequence data for this patent did not form part of the printed specification, but CC was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 558 AA;
Query Match 54.7%; Score 2755; DB 4; Length 558;
Best Local Similarity 93.3%; Pred. No. 1.le-240;
Matches 539; Conservative 1; Mismatches 4; Indels 34; Gaps 5;
QY 1 MSSGLWSQKVTSPYWEERIFYLLQECVTDKQTKLLKVPKSGISQYITQDRSVGHSRI 60
DB 8 MSSGLWSQKVTSPYWEERIFYLLQECVTDKQTKLLKVPKSGISQYITQDRSVGHSXI 67
QY 61 PSAKGGKQIGIKILEQPHAVLPVDE-DVVEINEKETELLALITNCEERESLFKNRRLS 119
DB 68 PSAKGGKQIGIKILEQPHAVLPVDEKDVVEINEKETELLALITNCEERESLFKNRRLS 127
QY 120 KGLQIDVGCPCVKQLRSGBEKEFGVVRFRGPIIAERTVSGIFFGVLELLEBGRGQFTDGV 179
DB 128 KGLQIDVGCPCVKQLRSGBEKEFGVVRFRGPIIAERTVSGIFFGVLELLEBGRGQFTDGV 187
QY 180 YQKQLFQCDDECG-FVALDKLELIEDDDTALSDYAGPGDTMQVELPPLPPIEINRSVSLKG 238
DB 188 YQKQLFQCDDECGFVALDKLELIEDDDTALSDYAGPGDTMQVELPPLPPIEINRSVSLK- 246
QY 239 GETIESGTWIFCDVLPKESLGYFVGVDMDNPIGNWDGRFDGV-LCSFACVESTILLHIN 297
DB 247 -----DNPIGNWDGRFDGVQVLCFACVESTILLHIN 277
QY 298 DIIPESVTOERRPPKLAFMRSRGVGDKGSSHNPKATGSTSDPCNR-RSELYFTLNGSSV 356
DB 278 DIIPESVTOERRPPKLAFMRSRGVGDKGSSHNPKATGSTSDPCNRNRESELFTLNGSSV 337
QY 357 DSQPSQSKNNTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPPVNSLTNRRFHSLPFSL 416
DB 338 DSQPSQSKNNTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPPVNSLTNRRFHSLPFSL 397
QY 417 TKMPNTNGSIHSPSLSLSAQSVMEELNTAPVQSPPLAMPBGSHGLEVGSIAEVKENPP 476
DB 398 TKMPNTNGSIHSPSLSLSAQSVMEELNTAPVQSPPLAMPBGSHGLEVGSIAEVKENPP 457
QY 477 FYGVIRWIGOPPGLENEVLAGELEDEACAGCTDGTFRGTRYFTCALKKALFVKLKS CRPDS 536
DB 458 FYGVIRWIGOPPGLENEVLAGELEDEACAGCTDGTFRGTRYFTCALKKALFVKLKS CRPDS 517
QY 537 RFASLQPVSNQIERCNSLAFGGYLSVEVEENTPPKMEK 574
DB 518 RFASLQPVSNQIERCNSLAFGGYLSVEVEENTPPSSEE 555
RESULT 14
AAB95719
ID AAB95719 standard; protein; 476 AA.
XX
AC AAB95719;
XX
XX 26-JUN-2001 (first entry)
DT
XX
XX Human protein sequence SEQ ID NO:18587.
DE
XX
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
XX Homo sapiens.
XX
XX EPI074617-A2.
XX
XX 07-FEB-2001.
XX
XX 28-JUL-2000; 2000EP-00116126.
PF


```

XX 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
XX
XX Claim 8; SEQ ID NO 18587; 2537pp + Sequence Listing; English.
XX
XX The present invention describes primer sets for synthesizing 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dt primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
XX Sequence 476 AA;
XX
Query Match 46.5%; Score 2343; DB 4; Length 476;
Best Local Similarity 98.1%; Pred. No. 2.3e-203;
Matches 454; Conservative 2; Mismatches 3; Indels 4; Gaps 3;
QY 221 MQVELPPLLEINRSVSLKGGTIESGTVIFCDVLPKGESLGTFVGVDMNDNPIGNWDRFGD 280
DB 1 MQVELPPLLEINRSVSLKGGTIESGTVIFCDVLPKGESLGTFVGVDMNDNPIGNWDRFGD 60
QY 281 V-LCSFACVESTILLHINDIIPESVTQERRPKPLAFMSRGVGDGKSSSHNKPATGSTSD 339
DB 61 VOLCSFACVESTILLHINDIIPESVTQERRPKPLAFMSRGVGDGKSSSHNKPATGSTSD 120
QY 340 PGNR-RSELYTLNGSSVDSQPSKSNWTYIDEVAEDPAKSLTETSTDFDRSSPPLQPP 398
DB 121 PGNRRSELYTLNGSSVDSQPSKSNWTYIDEVAEDPAKSLTETSTDFDRSSPPLQPP 180
QY 399 PVSNTLTENRPSLTPSLTKMPNTNGSIGHSPLSLSAQSWMEELNAPVQESPLAMP 458
DB 181 PVSNTLTENRPSLTPSLTKMPNTNGSIGHSPLSLSAQSWMEELNAPVQESPLAMP 240
QY 459 NSHGLEVGSLAEVKENPPFYGVIRWTQPPGLNEVLGLAGLEDECACTDGTFRGTRYFT 518
DB 241 NSHGLEVGSLAEVKENPPFYGVIRWTQPPGLNEVLGLAGLEDECACTDGTFRGTRYFT 300
QY 519 CALKALFKVLKSCRDPDSFASLQPVSNQIERCNSLAFGGYLSVVVEENTPPKMEKEGLE 578
DB 301 CALKALFKVLKSCRDPDSFASLQPVSNQIERCNSLAFGGYLSVVVEENTPPKMEKEGLE 360

```

```

QY 579 IMIGKKKGIQGHVNSCYLDSTLFCFLPAPSSVLDTVLLRPKPKNDVVEYSETQELLRTETV 638
DB 361 IMIGKKKGIQGHVNSCYLDSTLFCFLPAPSSVLDTVLLRPKPKNDVVEYSETQELLRTETV 420
QY 639 NPLRIYGVVCAATKIMKILKILEKVEAASGGFTSEKDPPEFLNI 681
DB 421 NPLRIYGVVCAATKIMKILKILEKVEAASGGFTSEKDPPEFLNI 461

RESULT 15
ABB89233
ID ABB89233 standard; protein; 261 AA.
XX AC ABB89233;
XX DT 24-MAY-2002 (first entry)
XX DE Human polypeptide SEQ ID NO 1609.
XX
XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
XX anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; anti-ulcer;
XX vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
XX cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
XX neurological disease; infection; human; secreted protein.
XX
XX Homo sapiens.
XX
XX WO200190304-A2.
XX
XX 29-NOV-2001.
XX
XX 18-MAY-2001; 2001WO-US016450.
XX
XX 19-MAY-2000; 2000US-0205515P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Birse CE, Rosen CA;
XX
XX WPI; 2002-122018/16.
XX
XX N-PSDB; ABL89642.
XX
XX Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
XX prevention of neural, immune system, muscular, reproductive,
XX gastrointestinal, pulmonary, cardiovascular, renal and proliferative
XX disorders.
XX
XX Claim 11; SEQ ID NO 1609; 2081pp + Sequence Listing; English.
XX
XX The invention relates to novel genes (ABL89449-ABL90853) and proteins
XX (ABB9040-ABB90444) useful for preventing, treating or ameliorating
XX medical conditions e.g. by protein or gene therapy. The genes are
XX isolated from a range of human tissues disclosed in the specification.
XX The nucleic acids, proteins, antibodies and (ant)agonists are useful in
XX the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
XX ovarian cancer and other cancers of the adrenal gland, bone marrow,
XX breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
XX disorders e.g. Addison's disease, allergies, autoimmune haemolytic
XX anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
XX multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
XX cardiovascular disorders such as myocardial ischaemias; (d) wound healing
XX ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
XX infectious diseases such as viral, bacterial, fungal and parasitic
XX infections. Note: The sequence data for this patent did not form part of
XX the printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 261 AA;
XX
Query Match 25.3%; Score 1275; DB 5; Length 261;
Best Local Similarity 98.7%; Pred. No. 1.1e-106;
Matches 234; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

Qy	653	MKLRKILEKVEAASGFTSEKDPPEEFNLILFHHILRVEPLLKIRSAGQKVQDCYFYQIEM	712
Db	1	MKLRKILEKVEAASGFTSEKDPPEEFNLILFHHILRVEPLLKIRSAGQKVQDCYFYQIEM	60
Qy	713	EXNEKVGVPITIOQLLEWSFINSNLKFAEAPSCLIIOMPREGKDFKLFKKIFPSLELNITD	772
Db	61	EXNEKVGVPITIOQLLEWSFINSNLKFAEAPSCLIIOMPREGKDFKLFKKIFPSLELNITD	120
Qy	773	LLEDTPROCRICGGLAMYECRECYDDPDISAGKIQFCKTCNTQVHLHPKRLNHNKYNPVS	832
Db	121	LLEDTPROCRICGGLAMYECRECYDDPDISAGKIQFCKTCNTQVHLHPKRLNHNKYNPVS	180
Qy	833	LPKDLPDWDRHGCIPQNMELFAVLCTSHYVAFVYKGDSDAWLFFDSMADRDG	889
Db	181	LPKDLPDWDRHGCIPQNMELFAVLCTSHYVAFVYKGDSDAWLFFDSMADRDG	237

Search completed: April 18, 2005, 15:23:39
 Job time : 79 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 18, 2005, 15:32:50 ; Search time 6177 Seconds
(without alignments)
5847.983 Million cell updates/sec

Title: US-09-671-687A-3

Perfect score: 5034

Sequence: 1 MSSGLWSQKVTSPWEEERI.....RLLCDAYMCYQSPMSLYK 949

Scoring table:

	BLOSUM62	Xgapop 10.0	Xgapext 0.5
Ygapop 10.0		Ygapext 0.5	
Fgapop 6.0		Fgapext 7.0	
Delop 6.0		Delext 7.0	

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

```
-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgn2.1/USPTO spool_p/US09671687/runat 18042005 115018 17139/app_query.fasta_1.1095
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOPCL=0 -LOPEXT=0
-UNITS=bts -START=1 -END=1 -MATRIX=blomumc2 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09671687 @CCN 1 1 5180 @runat 18042005 115018 17139 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database :

```
EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	ID	Description
1	4494	89.3	3496	3 AF161542	Homo sapi
2	4313	85.7	2862	9 AY406374	Homo sapi
3	4102	81.5	2730	9 AY406376	Mus muscu
4	3768	74.9	2862	9 AY406375	Pan trogl
5	2952	58.6	2220	3 AK039054	Mus muscu
6	2769	55.0	3137	3 BC028885	Mus muscu
7	1572.5	31.2	1063	5 BQ433523	AGENCOURT
8	1570	31.2	1039	4 BM457960	AGENCOURT
9	1459.5	29.0	1478	3 AK042764	Mus muscu

10	1435	28.5	938	4	BI666276	603290787
11	1426	28.3	2218	3	AK048183	Mus muscu
12	1416.5	28.1	1133	4	BM480127	AGENCOURT
13	1384	27.5	812	6	CA512526	UI-R-FJO-
14	1363	27.1	1892	3	BC024596	Mus muscu
15	1340	26.6	837	4	BG867631	602787414
16	1302	25.9	816	5	BQ442079	UI-N-EXO-
17	1288	25.6	721	5	BP162259	BP162259
18	1286.5	25.6	889	5	BUS05487	AGENCOURT
19	1280	25.4	783	4	BM016881	603643383
20	1268.5	25.2	715	5	CN157036	944487 MA
21	1267.5	25.2	840	5	BP144628	BP144628
22	1262	25.1	837	1	AU137267	AU137267
23	1235	24.5	726	7	CR772310	DKP2p468L
24	1233	24.5	717	7	CF750356	UI-M-HJO-
25	1215	24.1	758	1	AU122742	AU122742
26	1202.5	23.9	791	7	CN461343	UI-M-HNO-
27	1198	23.8	743	6	CD351332	UI-M-GIO-
28	1170	23.2	677	6	CD626856	56076837J
29	1165	23.1	648	1	AL707441	DKP2p686M
30	1165	23.1	697	4	BM724143	UI-E-EOI-
31	1160.5	23.1	781	7	CK636353	UI-M-HNO-
32	1158	23.0	826	7	CK597829	AGENCOURT
33	1152	22.9	708	6	CB466941	732655 MA
34	1152	22.9	794	2	BF141863	601791346
35	1151	22.9	770	4	BI906770	603064511
36	1139.5	22.6	785	6	CB235676	AGENCOURT
37	1128	22.4	906	5	BU461394	603774393
38	1127.5	22.4	907	5	BQ438227	AGENCOURT
39	1118	22.2	702	2	BB648023	BB648023
40	1117.5	22.2	755	7	CO404877	AGENCOURT
41	1107.5	22.0	723	6	CD559213	AGENCOURT
42	1105.5	22.0	1134	5	BM904688	AGENCOURT
43	1102	21.9	637	2	AW985722	uf85b05.Y
44	1094	21.7	639	7	CN704295	E0485H10-
45	1092.5	21.7	1365	3	AY383658	Rattus no

ALIGNMENTS

RESULT 1	AF161542	3496 bp	mrna	linear	HTC 22-MAY-2001
LOCUS	AF161542	Homo sapiens	HSPC057	mrna	complete cds.
DEFINITION	AF161542	Homo sapiens	HSPC057	mrna	complete cds.
ACCESSION	AF161542.1	GI:6841351			
VERSION	AF161542.1	GI:6841351			
KEYWORDS	HTC.				
SOURCE	Homo sapiens				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 3496)				
AUTHORS	Zhang, Q.H., Ye, M., Wu, X.Y., Ren, S.X., Zhao, M., Zhao, C.J., Fu, G., Shen, Y., Fan, H.Y., Lu, G., Zhong, M., Xu, X.R., Han, Z.G., Zhang, J.M., Tao, J., Huang, Q.H., Zhou, J., Hu, G.X., Gu, J., Chen, S.J., and Chen, Z.				
TITLE	Cloning and functional analysis of cDNAs with open reading frames for 300 previously undefined genes expressed in CD34+ hematopoietic stem/progenitor cells				
JOURNAL	Genome Res. 10 (10), 1546-1560 (2000)				
MEDLINE	20499367				
PUBMED	11042152				
REFERENCE	2 (bases 1 to 3496)				
AUTHORS	Zhang, Q.H., Ye, M., Zhou, J., Shen, Y., Wu, X.Y., Guan, Z.Q., Wang, L., Fan, H.Y., Mao, Y.F., Dai, M., Huang, Q.H., Chen, S.J., and Chen, Z.				
TITLE	Human full length cDNA cloned from cd34+ stem cells				
JOURNAL	Unpublished				
REFERENCE	3 (bases 1 to 3496)				
AUTHORS	Zhang, Q.H., Ye, M., Zhou, J., Shen, Y., Wu, X.Y., Guan, Z.Q., Wang, L., Fan, H.Y., Mao, Y.F., Dai, M., Huang, Q.H., Chen, S.J., and Chen, Z.				
TITLE	Direct Submission				
JOURNAL	Submitted (14-MAY-1999) Shanghai Institute of Hematology, Shanghai Second Medical University, Rui-Jin Hospital, 197 Rui-Jin Road II, Shanghai 200025, P. R. China				

FEATURES

Location/Qualifiers
 1..3496
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CBFAOB03"
 /tissue_type="blood"
 1517..3064
 CDS
 /codon_start=1
 /product="HSPC057"
 /protein_id="AAF29029.1"
 /db_xref="GI:6841352"
 /translation="MENTINGSIHSPLSASQVMEELNTAPVQESPPILAMPNGSHG
 LEVGSIAEYKRNPPFYGIVRWIGQPFGLNEVLAGELEDEACGCTDGTFRGTRYFTCA
 LKALFVKLSRDPDRFASLOPVSNOIERCNISLAFGYLSVEVEENTPPKMEKEGLE
 INMKKGIQGHVNSCYLSTLPCLFAPSSVLDITLLRPKKNDVYYSSETOLLRTE
 IYNPLRYGVCACTIMKILKILEVEAASGFTSEKDPDEFNLILFHHLLRVEPLLK
 IRSAGKQDCCFYQIFMSEKNEKVGPTTQQLLEWFSFNSNLKFAEAPSCLLIQMRPF
 KDFKLFPKLI FPSLELIQYILKLPDSAGYVEGLQMSVENATTITFSAGKIQKQFC
 KTCNTQHLHPRLNHYNPVLSPKDLPDWDRHGCI PCQNMELFVLCIETSHYVAF
 VKYKDDSAWLFPDSMDRGGQNGFNIPQVTPCEVGBYLKMSLEDLHSLDSRRIPR
 LCTKTAL"

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 3496
 Score: 4494.00 Matches: 917
 Percent Similarity: 95.33% Conservative: 2
 Best Local Similarity: 95.12% Mismatches: 30
 Query Match: 89.27% Indels: 22
 DB: 3 Gaps: 4

US-09-671-687A-3 (1-949) x AF161542 (1-3496)

QY	1	MetSerSerGlyLeuTrpSerGlnGlu-LysValThrSerProTyrTrp-GluGluArgI	20
DB	248	ATGAGTTTCAGGCTATGGAGCCCAAGAAAAAGTCACTTCCCTACTCTGGGGAAGAGCGGA	307
QY	20	lePheTyrLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLeuL	40
DB	308	TTTTTTTACTTGCTTCTTCAAGAAATGACGCGTTACAGCAACAAACAAACAAAGCTCCTTA	367
QY	40	ysValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgI	60
DB	368	AAGTACCGAA-GGAGTATAGGACAGTATATTCAAGATCGTCT-GTGGGCGATTCAAGGA	425
QY	60	leProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisA	80
DB	426	TTCTCTCTGCAAAAGGCAG-AAAAATCAGATTGGATTAAAAATTTCTAGAGCAACCTCATG	484
QY	80	laValLeuPheValAsp-Glu---AspValValGluIleAsnGlnLysPheThrGluLeu	98
DB	485	CAGTTCTCTTTGTGTGTAAGAAAGGATGTTGTAGAGATGAATGAAAGATTACAGAGATTT	544
QY	99	LeuLeu---AlaIleThrAsnCys-GluGluArgPheSerLeuPheLysAsnArgAsnAr	117
DB	545	ACTTTTGGGCAATACCCTAATTTGGAGAGAGGTTTCAGGCTGTTTAAAAACAGAAACAG	604
QY	117	gLeuSerLysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGI	137
DB	605	ACTAAGTAAAGGCTCCCAATAGACGTGGGCTGCTCTGTGAAAGTACAGCTGAGATCTGG	664
QY	137	yGluGluLysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrVa	157
DB	665	GGAAAGAAAAATTTCTCGAGTTGTACGCTTTACAGAGGACCCCTGTTTAGCAGAGAGGACAGT	724
QY	157	lSerGlyIlePhePheGlyValGluLeuLeuGluGluArgGlyArgGlyGlnGlyPheThrAs	177
DB	725	CTCCGGAAATATCTTTTCGAGTTGAATTTGTG-GAAGAAGGTCGTGGTCAAGGTTTCACTGA	783
QY	177	pGlyValTyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAl	196
DB	784	CGGGGTGTACCAAGGGAAAACAGCTTTTTCAGTTGTGATGAATATTGTGCGCGTGTGTTGTC	843
QY	196	alLeuAspLysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlu	216
DB	844	ATTGGCAAGCTAGAACTCATAGAGATGATGACACTGCATTTGGAAAGTATTACGCAGG	903
QY	216	yProGlyAspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLe	236
DB	904	TCCTGGGGACACAATGAGTCGAACTTCTCTCTTTGGAAATAAATCCAGAGTTCTTT	963
QY	236	uLysGlyGlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLY	256
DB	964	GAA-GGTTGAGAAACAATAGATCTGGACAGTTATATTCTGTGATGTTTTTGCAGGAAA	1022
QY	256	sGluSerLeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGI	276
DB	1023	AGAAAGCTTAGGATATTTTGGTGGTGGACATGGATAACCTTATTGGCAACTGGGATGG	1082
QY	276	yArgPheAspGlyValLeuCysSer-PheAlaCysValGluSerThrIleLeuLeuHisI	296
DB	1083	AAGATTTGATGGAGCGCAGCTTGTAGTTTGCGTGTGTGAAAGTCCAATTTCTATTGCACA	1142
QY	296	leAsnAspIleIleProGluSerValThrGlnGluArgArgProProLysLeuAlaPheM	316
DB	1143	TCAATGATATCATCCAGAGAGTGTGACGCAAGAAAGAGGCGCTTCCAAACTTGCCTTTA	1202
QY	316	etSerArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThrGlyS	336
DB	1203	TGTCAGAGAGTGTGGGGCAAAAGTTTCATC-AGTCATATAAACCACCAAGGCTACAGAT	1261
QY	336	erThrSerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerS	355
DB	1262	CTACCTTAGACCTCGGAAATAGAAAACAGATCTGAAATTTATATACCTTAAATGGGTCTT	1321
QY	355	erValAspSerGln-ProGlnSerLysSerLysAsnThrTrpTyr-IleAspGlu-ValA	374
DB	1322	CTGGTACCTTACAACTCCCAATCCAAATCAAAATCAAAATCCATGGTACCATTGATGAAAGTTG	1381
QY	374	laGluAspProAlaLysSerLeuThrGluIleSerThr-AspPheAspArgSerSerPro	393
DB	1382	CATAAAGACCTTGAAGAACTCTTACAGAGATATCTACCCGACCTTTGCCCGGTCTTCACCA	1441
QY	394	ProLeuGlnProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuPro	413
DB	1442	CCACTCCAGGCTCCTCTGTGAACTCACTGACCACCGAGAACAGATTCCACTCTTTACCA	1501
QY	414	PheSerLeuThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeu	433
DB	1502	TTCAGTCTTCACCAAGATGCCCAATACCAATGGAAATATTGGCCACAGTCCACTTTCTCTG	1561
QY	434	SerAlaGlnSerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeu	453
DB	1562	TCAGCCAGCTCTGTATATGGAGAGCTTAACACATGCACCCGTCGAAGAGTCCACCCCTTG	1621
QY	454	AlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGlu	473
DB	1622	GCCATGCTCTCTGGAACTCACATGGTCTAGAGTGGGCTCATTTGGTGAAGTTAAGGAG	1681
QY	474	AsnProProPheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluVal	493
DB	1682	AACCTCTCTTCTATGGGTAATCCGTTGGATCGGTGAGTCCAGCCACCCAGGACTGAATGAAGTG	1741
QY	494	LeuAlaGlyLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGly	513
DB	1742	CTCGCTGAGCTGGAACCTGGAAGATGAGTGTGCGGCTGTACGAGTGAACCTTCAGAGGC	1801
QY	514	ThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArg	533
DB	1802	ACTCGGTATTTTCACTGTGCCCTGAAGAAAGGCGCTGTTTGTGAAACTGAAGAGCTGCAGG	1861
QY	534	ProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSer	553
DB	1862	CCTGACTCTAGGTTTGATCATTTGTCAGCCGCGTTTCCAATCAGATTGAGCGCTGTAACCTCT	1921

554 LeuAlaPheGlyGlyTyrLeuSerGluValValGluAsnThrProProLysMetGlu 573
1922 TTAGCAATTTGGAGGCTACTTAAGTAGAGTAGAGAAATACTCCACCAAAATGGAA 1981
574 LysGluGlyLeuGluMetIleGlyLysLysGlyGlyIleGlnGlyHisTyrAsnSer 593
1982 AAAGAAGGCTTGGAGATAATGATTGGGAAGAAGAGGCATCCAGGCTCATTACAAATCT 2041
594 CysTyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrVal 613
2042 TGTACTTAGACTCAACCTTATTCTGCTTATTGCTTTAGTCTTGTCTGGACACTGTG 2101
614 LeuLeuArgProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeu 633
2102 TTACTTAGACCCCAAGAAAGAAAGACGATGATAGATATTATAGTGAACCCCAAGAGCTACTG 2161
634 ArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMet 653
2162 AGGACAGAAATTTGTTATCTCTGAGATATATGATATGTTGTGCCACAAAATATG 2221
654 LysLeuArgLysIleLeuGluLysValGluAlaSerGlyPheThrSerGluGluLys 673
2222 AAACCTGAGGAATACTTGAAGAGGTGGAGGCTGCATCAGGATTTACCTCTGAAGAAAA 2281
674 AspProGluGluPheLeuAsnIleLeuPheHisHisIleLeuArgValGluProLeu 693
2282 GATCTCGAGGAATCTTGAATATTTCTGTTTCATCATATTTTAAAGGGTAGAACCTTTGCTA 2341
694 LysIleArgSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGlu 713
2342 AAATTAAGATCAGCAGCTCAAAGGTACAAAGATTGTACTTCTATCNAATTTTATGGAA 2401
714 LysAsnGluLysValGlyValProThrIleGlnGlnLeuLeuGluTrpSerPheIleAsn 733
2402 AAAAATGAGAAAGTGGCGTTCCCACAATTCAGCAGTTGTTAGATGCTCTTTATCAAC 2461
734 SerAsnLeuLysPheAlaGluAlaProSerCysIleuIleleGlnMetProArgPheGly 753
2462 AGTAACCTGAAATTTGAGAGGACCACCATCATGTCTGATTATTCAGATGCCCTCGATTGGA 2521
754 LysAspPheLysLeuPheLysIlePheProSerLeuLeuGluLeuAsnIleThrAspLeu 773
2522 AAAGACTTTAACTATTATAAAAAATTTTTCCTTCTCTGGAAT-ATAATACAGATTTA 2580
774 LeuGluAspThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArg 793
2581 CTTGAAGACACTCCGACAGTCCCGATATGTGGAGGCTTGCATGATGATGAGTGAGA 2640
794 GluCysTyrAspAspProAspIle-SerAlaGlyLysIleLysGlnPheCysLysThrCy 813
2641 GAATGCTACGACGATCCGACATTCCTCAGCTGGAAAAATCAAGCAGTTTGTAAACCTG 2700
813 sAsnThrGlnValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLe 833
2701 CAACACTCAAGTCCACCTTCATCCGAAGAGGCTGAATCAATAATAATATACCCAGTGCTCACT 2760
833 uProLysAspLeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGl 853
2761 TCCCAAGACTTACCCGACTGGAGCTGGAGACACCGCTGCATCCCTTGCAGATATGGA 2820
853 uLeuPheAlaValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLy 873
2821 GTTATTGCTGTTCTCTGCATAGAAACAAGCCACTATGTTGCTTTTGTGAAGTATGGAA 2880
873 sAspAspSerAlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGl 893
2881 GGACGATTCTCGCTGGCTCTCTTTTGACAGCATGGCCGATCGGGATGGTGGTCAAGATGG 2940
893 yPheAsnIleProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLe 913
2941 CTTCAACATTCCTCAAGTACCCCATCCCATCCAGAGTAGGAGAGTGTGAAGATGTCTCT 3000
913 uGluAspLeuHisSerLeuAspSerArgArgIle-GlnGlyCysAlaArgArgLeuLeuC 933

3001 GGAAGACCTGCATTCCTTGGACTCCAGAGATCCCAAGGCTGTGCAGAGACTGCTTT 3060
933 ysAspAlaTyrMetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
3061 GTGATGCATATATGTGATGTACAGAGTCCAAACATGAGTTTGTACAAA 3110
RESULT 2
AY406374 2862 bp DNA linear GSS 15-DEC-2003
LOCUS Homo sapiens CYLD gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY406374
VERSION AY406374.1 GI:39762348
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2862)
AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarimal,A.,
Todd,M.A., Tanenbaum,D.M., Civeello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene tritos
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 2862)
AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarimal,A.,
Todd,M.A., Tanenbaum,D.M., Civeello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
Location/Qualifiers
source 1..2862
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
gene <1..>2862
/genes="CYLD"
/locus_tag="HCM2536"
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 2862
Score: 4313.00 Matches: 839
Percent Similarity: 88.04% Conservative: 0
Best Local Similarity: 88.04% Mismatches: 110
Query Match: 85.68% Indels: 4
DB: 9 Gaps: 4
US-09-671-687A-3 (1-949) x AY406374 (1-2862)
Qy 1 MetSerSerGlyLeuTyrProSerGlnGluLysValThrSerProTyrTrpGluGluArgile 20
Db 1 ATGAGTTTCAGGCTTATGGAGCCCAAGAAAAAGTCACTTCACCCCTACTGGGAAGCGGATT 60
Qy 21 PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLeuLys 40
Db 61 TTTTACTTGTCTTCTCAAGATTCGAGCTTACAGACAAACAAACAAACAAACCTCTTAAA 120
Qy 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgile 60
Db 121 GTACCGAAGGGAAGTAGGACAGTATATTCAAGATCGTTCTCTGGGGCAATCAAGGATT 180
Qy 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla 80
Db 181 CCTTCTCAAAAGCAAGAAAAATTCAGATTGGATTAAAAATTTCTAGACCAACCTCATGCA 240

```
QY 81 ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu 99
DB 241 GTTCTCTTTGTCATGAAAAGGATGTTGTAGAGATAAATGAAAGATTTCACAGAGTTACTT 300
QY 100 LeuAlaIleThrAsnGlyValGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
DB 301 TTGGCAATTTACCAATTTGTGAGGAGAGGTTGAGGCTGTTTAAACAGAAACAGACTAAGT 360
QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
DB 361 AAAGGCTCCAAATAGACGTGGGCTGCTGTGTGAAGATACAGCTGAGATCTGGGGAAGAA 420
QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
DB 421 AAATTTCTCTGGAGTTTACGCTTCAGAGACCCCTGTTAGCAGAGAGGACAGTCTCCGGA 480
QY 160 IlePhePheGlyValGluLeuLeuGluGluGlyArgGlyGlnGlyPheThrAspGlyVal 179
DB 481 ATATTTCTTTGGAGTTGAATTTGCTGGAGAAGGTCGTGTCAGGTTTTCATGACGGGGTG 540
QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
DB 541 TACCAAGGGAACAGCTTTTTCAGTGTGATGATGAAGATTGGCGTGTGTTGTTGCTTTGGAC 600
QY 199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
DB 601 AAGCTAGAACTCATAGACATGATGACATGCTGATTTGGAAGTGATTACGCGAGGCTCTGGG 660
QY 219 AspThrMetGlnValGlu---LeuProProLeuGluIleAsnSerArgValSerLeuLys 237
DB 661 GACACAATGCGGTGCAAAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 720
QY 238 GlyGlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGlu 257
DB 721 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 780
QY 258 SerLeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArg 277
DB 781 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 840
QY 278 PheAspGlyValLeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
DB 841 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 900
QY 298 AspIleIleProGluSerValThrGlnGluArgProProLysLeuAlaPheMetSer 317
DB 901 NNNNNNNNNNNNAGAGTGTGCGCAGAGAAAGGAGGCTCCCAAACTTGCTTTATGTCA 960
QY 318 ArgGlyValGlyAspLysGlySerSerSerHisAsnLysProLysAlaThrGlySerThr 337
DB 961 AGAGGTGTTGGGACAAAGGTTTCATCCAGTCAATAAACAAGGCTACAGGATCTACC 1020
QY 338 SerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVal 356
DB 1021 TCAGACCTCGAAATAGAAACAGATCTGAATTTATTTATACCTTAAATGGGTCTCTGTT 1080
QY 357 AspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGluValAlaGluAsp 376
DB 1081 GACTCAACCAACCAATCCAAATCAAAATAATACATGTTGATGATGATGATGATGATGATGAT 1140
QY 377 ProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProProLeuGln 396
DB 1141 CTTGCAAAATCTCTTACAGAGATATCTACAGACTTTTGACCGTCTTCCACCACTCCAG 1200
QY 397 ProProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeu 416
DB 1201 CCTCTCTCTGAACTCACTGACCCAGACAGAAACAGATTTCCACTCTTTTACCATTCCAGTCTC 1260
QY 417 ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGln 436
DB 1261 ACCAAGATGCCCAATCAACATGGAAGTATTGGCCACAGTCCACTTTCTCTGTCAGCCACG 1320
```

```
QY 437 SerValMetGluLeuLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPro 456
DB 1321 TCTGTAAATGGAAGAGCTAAACACTGCACCGCTCAAGAGAGTCCACCTTTGGCCATGCCT 1380
QY 457 ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro 476
DB 1381 CTTGGGAACCTCATGCTCTAGAACTGGGCTCATTTGGCTGAAAGTTAAGGAGAACCTCTCT 1440
QY 477 PheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGly 496
DB 1441 TTCTATGGGGTAATCCGTTGGATCGGTACGCCACAGGACTGAATGAAGTCTCGCTGGA 1500
QY 497 LeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr 516
DB 1501 CTGGAACCTGGAAGATGATGTGCGAGGTGTACGATGGAACCTTCAGAGGCACTCGGTAT 1560
QY 517 PheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSer 536
DB 1561 TTCACCTGTGCCCTGAAGAAGGCGCTGTTTGTGAACTGAAGAGCTGCAGGCTGACTCT 1620
QY 537 ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe 556
DB 1621 AGGTTTGCATCATTTGCAGCGGTTTCCAAATCAGATTGAGCGCTGTAACCTCTTTAGCATTT 1680
QY 557 GlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluGly 576
DB 1681 GGAAGGCTACTTAAGTGAAGTAGTAGAAGAAATACTCCACCAAAATGGAAGAAAGGC 1740
QY 577 LeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu 596
DB 1741 TTGGAGATTAATGATTGGGAAGAAAGGCGCATCCAGGCTCATTAACAATCTTTGTACTTTA 1800
QY 597 AspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArg 616
DB 1801 GACTCAACCTTATCTGCTTATTTGCTTTAGTTCTGTTCTGGACACTGTGTACTTAGA 1860
QY 617 ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu 636
DB 1861 CCCAAAGAAAAGAACGATGTAGAAATATATAGTGAACCCCAAGAGCTACTGAGGACAGAA 1920
QY 637 IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg 656
DB 1921 ATTGTTAAATCCTCGAGAATNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1980
QY 657 LysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGlu 676
DB 1981 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2040
QY 677 GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArg 696
DB 2041 GAATTCCTTGAATATCTGTTTTCATCATATTTTAAAGGCTGAACCTTTGCTAAAAATAAGA 2100
QY 697 SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu 716
DB 2101 TCAGCAGGTCAAAAGGTACAAAGATTGTTACTTCTATCAAAATTTTATGGAAAAAATGAG 2160
QY 717 LysValGlyValProThrIleGlnLeuLeuGluTrpSerPheIleAsnSerAsnLeu 736
DB 2161 AAAGTTGGCGCTCCCAATTCAGCAGTGTGTTAGATGGTCTTTTATCAACAGTAACCTG 2220
QY 737 LysPheAlaGluAlaProSerCysLeuIleGlnMetProArgPheGlyLysAspPhe 756
DB 2221 AAATTTGCAGAGGACCATCATGCTGATTATTTCAGATGCCTCGATTTGGAAAAAGACTTT 2280
QY 757 LysLeuPheLysIlePhePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAsp 776
DB 2281 AAACTATTTAAAAAATTTTCTCTCTCGAAATTAATAATAACAGATTTTACTTTGAAGAC 2340
QY 777 ThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyr 796
DB 2341 ACTCCCAAGACAGTCCCGATATGTGGAGGGCTTCCAATGTATGATGTAGAGATGCTTAC 2400
QY 797 AspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGln 816
```


gene trios
 Science 302 (5652), 1960-1963 (2003)
 14671302
 2 (bases 1 to 2862)
 Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejaritwal, A.,
 Todd, M.A., Tanenbaum, D.M., Civallo, D.R., Lu, F., Murphy, B.,
 Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
 Adams, M.D., and Cargill, M.
 Direct Submission
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 This sequence was made by sequencing genomic exons and ordering
 them based on alignment.

FEATURES
 source
 1..2862
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"
 <1..>2862
 /gene="CYLD"
 /locus_tag="HCW253c"

ORIGIN
 Alignment Scores:
 Pred. No.: 0 Length: 2862
 Score: 3768.00 Matches: 751
 Percent Similarity: 78.80% Conservative: 0
 Best Local Similarity: 78.80% Mismatches: 198
 Query Match: 74.85% Indels: 4
 DB: 9 Gaps: 4

US-09-671-687A-3 (1-949) x AY406375 (1-2862)

QY 1 MetSerSerGlyLeuTyrThrSerGlnGluLysValThrSerProTyrTyrGluGluArgIle 20
 DB 1 ATGAGTTTCTGCTTATGGAGCCAGAAAAGTCTCCTTCCCTTCTGCGGAGCGGATT 60
 QY 21 PheTyrLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLeuLys 40
 DB 61 TTTTACTTCTTCTTCAAGATGTCAGCGTTACAGACAAACAAACAAAGCTCTTAAA 120
 QY 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60
 DB 121 GTACCGAAGGGAAGTATAGGACAGTATATTCAGATCGTCTGCGGGCAATTCAGGATT 180
 QY 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGlnProHisAla 80
 DB 181 CCTTCTGCAAAAGGCAAGAAATCATGATTGGATTAAATTTCTAGACAACTCATGCA 240
 QY 81 ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu 99
 DB 241 GTTCTCTTTGTGATGAAAAGGATGTTGTAGAGATAAATGAAAGTTTCACAGATTACTT 300
 QY 100 LeuAlaIleThrAsnGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
 DB 301 TTGGCAATATACCAATNNNNNAGAGTTTCAGCTGTTTAAAAACAGAAACAGACTAAGT 360
 QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
 DB 361 AAANNNTTCCAAATAGNN 420
 QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
 DB 421 NNN 480
 QY 160 IlePhePheGlyValGluLeuLeuGluGlyArgGlyGlnGlyPheThrAspGlyVal 179
 DB 481 NNN 540
 QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
 DB 541 TACCAAGGGAACAGCTTTTTCAGTGTGATGAAGATTGTGGCGTGTGTTTGTGATGGAC 600

QY 199 LysLeuGluLeuIleGluAspAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
 DB 601 AAGCTAGAACTCATAGAAAGATGATGACACATGTCATTTGAAATGATGATGCGAGGCTCTGGG 660
 QY 219 AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly 238
 DB 661 GACACAATGACGCTCGAANN 720
 QY 239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
 DB 721 NNNNNACATAGAACTCGAACNNNTATATTTCTGTGATGTTNNNNNNNGAAAGAAAGC 780
 QY 259 LeuGlyTyrPheValGly---ValAspMetAspAsnProIleGlyAsnTyrAspGlyArg 277
 DB 781 TTAGGATATTTTGTGTGNN 840
 QY 278 PheAspGlyValLeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
 DB 841 NNN 900
 QY 298 AspIleIleProGluSerValThrGlnGluArgProProLysLeuAlaPheMetSer 317
 DB 901 NNNNNNNNNNNNNNTGTGACGCAAGAAAGGAGGCTCCCAAACTTGCCTTTATGTCA 960
 QY 318 ArgGlyValGlyAspLysGlySerSerSerHisAsnLysProLysAlaThrGlySerThr 337
 DB 961 AGAGGTGTTGGGGACAAAGGTTTCATCCAGTCATAATAAACCAAGGTTACAGGATCTACC 1020
 QY 338 SerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVal 356
 DB 1021 TCAGACCTGGAAATAGAAACAGATCTGAATATTTTATACCTTAAATGGGTCTTCTGTT 1080
 QY 357 AspSerGlnProGlnSerLysSerLysAsnThrTyrIleAspGluValAlaGluAsp 376
 DB 1081 GACTCACACCACAATCCAAATCAAAAATACATGGTACATTTGATGAAGTTGCAGAAAC 1140
 QY 377 ProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerProLeuGln 396
 DB 1141 CCTGCAAAATCTCTTACAGAGATATCTACAGACTTTTACCGTTCTTACCACCACTCCAG 1200
 QY 397 ProProValAsnSerLeuThrGluAsnArgPheHisSerLeuProPheSerLeu 416
 DB 1201 CCTCTCTCTGTGAACCTACCTGACCCAGAGAGAGATTCACACTCTTTACCATTGATCTC 1260
 QY 417 ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGln 436
 DB 1261 ACCAAGATGCCAATACCAATGGAAGTATTTGGCCACAGTCCACTTTCTGTCTAGCCAG 1320
 QY 437 SerValMetGluLeuLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPro 456
 DB 1321 TCTGTAATGGAAGAGCTAAACACTGCACCCCGTCCAGAGAGTCCACCCCTTGGCCATGCT 1380
 QY 457 ProGlyAsnSerHisGlyLeuValGlySerLeuAlaGluValLysGluAsnProPro 476
 DB 1381 CCTGGGAACCTCACATGCTTAGAAGTGGGCTCATTTGGCTGAAGTTAAGGAGAACCCCTCT 1440
 QY 477 PheTyrGlyValIleArgTyrIleGlyGlnProProGlyLeuAsnGluValLeuAlaGly 496
 DB 1441 TTTCTAGGAGTAATCTGATCGGTGCGTCAGCCACAGACTGAATGAAGTCTGCTCGA 1500
 QY 497 LeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr 516
 DB 1501 CTGGAACCTGGAAGATGAGTGTGAGGCTGTACGGATGGAACTTTCAGAGGCACCTCGGTAT 1560
 QY 517 PheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSer 536
 DB 1561 TTTCACTGTGCTTGAAGAGGCGCTGTTTGTGAACTGAAGAGCTGAGGAGCTGAGCTCT 1620
 QY 537 ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe 556
 DB 1621 AGTTTTCATCATTTGACCGCGTTTCCAATCAGATCGAGCGCTGTAATCTCTTTAGCATTT 1680
 QY 557 GlyGlyTyrLeuSerGluValValGluLeuAsnThrProProLysMetGluLysGluGly 576

Db 1681 GAGGGCTACTTAAGTGAAGTAGTAGAAGAAATACTCCACCAAAATGGAAAAGAGGC 1740
Qy 577 LeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu 596
Db 1741 TTGGAGATAATGATTGGGAAGAAGAACGATCCAGGGTCATTACAAATCTTGTTACTTA 1800
Qy 597 AspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArg 616
Db 1801 GACTCAACCTTATCTGCTATTGCTTTAGTTCTGTGACACTGTGTTACTTAGA 1860
Qy 617 ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu 636
Db 1861 CCCAAGAAAAGAACGATGAGAATATTATGAGAAACCCAGAGACTCTGAGGACAGAA 1920
Qy 637 IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg 656
Db 1921 ATTGTTAATCTCTGAGATNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1980
Qy 657 LysIleLeuGluLysValGluAlaLaSerGlyPheThrSerGluGluLysAspProGlu 676
Db 1981 NNN 2040
Qy 677 GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArg 696
Db 2041 NNN 2100
Qy 697 SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu 716
Db 2101 TCAGCAGGTCAAAAGGTACAAGATTGTTACTTCTATCAAAATTTTATGAGAAAATAAG 2160
Qy 717 LysValGlyValProThrIleGlnLeuLeuGluTyrSerPheIleAsnSerAsnLeu 736
Db 2161 AAAGTTGGCTGCCACAAATTCAGCAGTGTGTAGATGGTCTTTATCAACAGTAACCTG 2220
Qy 737 LysPheAlaGluAlaProSerCysLeuIleGlnMetProArgPheGlyLysAspPhe 756
Db 2221 AAATTTGCAGAGNN 2280
Qy 757 LysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuGluAsp 776
Db 2281 NNN 2340
Qy 777 ThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyr 796
Db 2341 NCTCCAGACAGTCCGGATATGTGAGGGCTTGCAATGATGATGAGTAGAGATGCTAC 2400
Qy 797 AspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGln 816
Db 2401 GACGATCCGGACATCTCAGCTGGAAATAATCAAGCAGTTTTGTAAACCTGCAACACTCA 2460
Qy 817 ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAsp 836
Db 2461 GTCCACCTTCATCCGAAGAGGCTGAATATAATAAACCCAGTGCATCTTCCCAAGAC 2520
Qy 837 LeuProAspThrAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAla 856
Db 2521 TTACCTGACTGGGACTGGAGACACGGCTGCATCCCTGCCAGAAATATGGAGTTATTGCT 2580
Qy 857 ValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSer 876
Db 2581 GTTCTCTGCATAGAAACGACCACTATGTTGCTTTGTGAAGTATGGGAAGCAGATTCC 2640
Qy 877 AlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlnAsnGlyPheAsnIle 896
Db 2641 GCCTGGCTCTCTTTGACACATGCGCGATGCGGATGGTGGTCCAGAACCGCTTCAACATT 2700
Qy 897 ProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLeu 916
Db 2701 CCTCAAGTACCCCATGCCAGAGTAGGAGAGTACTTGAAGATGCTCTGGAAGACCTG 2760
Qy 917 HisSerLeuAspSerArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyr 936

Db 2761 CATTCCTTGGACTCCAGGAATCCAGGCTGTGCAGAGACTGCTTTGTGATGCATAT 2820
Qy 937 MetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
Db 2821 ATGTGCATGTACCAGAGTCCAAACATGAGTTTGTACAAA 2859
RESULT 5
AK039054
LOCUS
DEFINITION
Mus musculus adult male hypothalamus cDNA, RIKEN full-length
enriched library, clone:A230090N11 product:CYLINDROMATOSIS (TURBAN
TUMOR SYNDROME) homolog (Homo sapiens), full insert sequence.
AK039054
ACCESSION AK039054.1 GI:26333008
VERSION
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM
REFERENCE 1
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuina, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujikawa, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6
AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, M.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Inotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome

Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.jp/

URL: http://fantom.gsc.riken.jp/

FEATURES

Location/Qualifiers

source

1..2220
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:A230090N11"
/db_xref="taxon:10090"
/clone="A230090N11"
/sex="male"
/tissue_type="hypothalamus"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
352..>2220

CDS

/note="unnamed protein product; CYLINDROMATOSIS (TURBAN
TUMOR SYNDROME) homolog [Homo sapiens] (SPTRI096SH0,
evidence: FASTV, 93.8%ID, 30.5%length, match=873)
putative"
/codon_start=1
/protein_id="BAC30222.1"
/db_xref="GI:26333009"
/translation="MSSGLWSQEKVTSPYWEIRFYLLQECSTVDKQKLVKPKVG
STGVYIDRSVGHRSVPSTGKNOIGLKEOPHVLVFEKDVVEINEKFTLELLA
TINCERLSLVRNRLSKGLQVDSVPKVLQSRGEKPGVRRGPIILARTVSG
IFPVELLEGGQGTGVYQSKQLFQCEDCGVFALDKLELIEDDNLLESDPAG
PGDTWQVEPPLEINRSVLSKVGESTGVIFCDVLPKGSGLGVYFVDMNDPIGNW
DRFDGVLQCSFASVSTILLHNDIIPALSDSVTOERRPPKLAFFMRGVGDKSSSH
NPKVTSPPQPSDGRNRSLSFVTLNGSSVDSQSKNPWIDEVAEDPAKSLTEMS
DFGHSPPQPSDGRNRSLSFVTLNGSSVDSQSKNPWIDEVAEDPAKSLTEMS
TPVQSPPLPISGNNAHGLEVSLEAKENPFYGVIRWIGQPPGLSDVLGLELDE
CAGCTDTPRGTRYFTALKALFVKLSRPSRFASISQVSNQICRNSLAFGGYL
SEVVEENTPPRMEKEGLEIMIGKKGIQGHYNSCYLDSTLFLCLFAPSSALDVTLLRP

ORIGIN

Alignment Scores:
Pred. No.: 8,73e-294 Length: 2220
Score: 2952.00 Matches: 575
Percent Similarity: 94.87% Conservative: 17
Best Local Similarity: 92.15% Mismatches: 24
Query Match: 58.64% Indels: 8
DB: 3 Gaps: 6

US-09-671-687A-3 (1-949) x AK039054 (1-2220)

QY 1 MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluGluArgIle 20
DB 352 ATGAGTTTACGGCTGTGGAGCCCAAGAGAAATGTTACTTCCCTCTGGGAGAACCGATT 411
QY 21 PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLys 40
DB 412 TTTTATCTGCTTCTCAAGATGCAGTGTAAACAGACAAACAACTCAGAGCTGCTGAAA 471
QY 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60
DB 472 GTACCCAAAGGGAGTAGTACAGACAGTACATCCAAAGCCGTCTCTGTGGGGCAATCAAGATT 531
QY 61 ProSerAlaLysGlyLysAsnGlnIleGlyLeuLysIleLeuGlnProHisAla 80
DB 532 CCTTCCACAAAGGCAAGAAATACAGATTGGATTAAATAATCTTGGGCAACCCGATGCA 591
QY 81 ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu 99

DB 592 GTTCTGTTTGTGATGAAAAGGATGTTGTAGAAATAAATGAAAATTCACAGAGTTACTG 651
QY 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
DB 652 TTGGCAATATACCAACTGTGAGGAGAGGCTCAGCCTATTATAGAAAACAGACTCCGACTAAGT 711
QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
DB 712 AAAGGGCTCCAGGTAGACGTGGGCGAGTCTGTGTGAAGTACAGCTGCGCATCTGGGGAAGAG 771
QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
DB 772 AAATTTCCAGGAGTTGTACGCTTCAGAGGACCTTTATTAGCGGAGAGGACGGTGTGGGG 831
QY 160 IlePhePheGlyValGluLeuLeuGluGluGlyArgGlyGlnGlyPheThrAspGlyVal 179
DB 832 ATTTTCTTTGGAGTAGAATATTATGGAAGAAGTCTGTGTCAGAGTTTCACGGATGGGTA 891
QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaIleAsp 198
DB 892 TATCAAGGGAAGCAGCTTTTCCAGTGTGATGAGGACTGTGCGCTGTGTTGTTGCAATTGGAC 951
QY 199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
DB 952 AAATCTAGAACTTATAGAAAGATGACAAATGGATTGGAAGTGAATTTTGCAGGCCGAGA 1011
QY 219 AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly 238
DB 1012 GATACAATGCAGGTGAACCTCCCTTTGGAATAAATCCACAGAGTTTCTTTGAAGGTT 1071
QY 239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
DB 1072 GGAGAAGTACAGAACTCTGGAACAGTAATATCTGTGATGTTTACCAGGAAAAAGAGAGT 1131
QY 259 LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe 278
DB 1132 CTAGGATATTTTGTGTGTGGACATGGAATACCTATTGGCAACTGGGATGGAGAGGTTT 1191
QY 279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
DB 1192 GATGGAGTACAGCTCTGTAGTTTTCGAAGTGTGTAAGTACAAATCTCTGCACATCAAT 1251
QY 298 AspIleIlePro-----GluSerValThrGlnGluArgArgProProLysLeuAla 314
DB 1252 GACATCATCCAGCTTTATCAGATAGGCTGACACAGGAAAGAGGCGCTCCCAACTTGC 1311
QY 315 PheMetSerArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThr 334
DB 1312 TTTATGTCAAGAGGTGTAGGTGACAAAGGTTTCATCTAGTACATAATAAACCAAGAGTTACA 1371
QY 335 GlySerThrSerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGly 353
DB 1372 GGATCTACCTCAGACCCCTGGAAGTAGAAACAGATCTGAAATATTATTATACCTTAATGG 1431
QY 354 SerSerValAspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGluVal 373
DB 1432 TCATCTGTTGACTCACAA---CAATCCAAGTCCAAAAATCCCATGCTGATGTAAGT 1488
QY 374 AlaGluAspProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerPro 393
DB 1489 GCAGAAAGACCCCTGCAAAAGTCACTTACAGAGATGCTTTCGGGACTTCGGACATTCATCTCT 1548
QY 394 ProLeuGlnProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuPro 413
DB 1549 CCACCGCAGCCTCTCTCCGTGAACCTCTGTGTAGCGAGAACAGATTCACCTCTCTTACC 1608
QY 414 PheSerLeuThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeu 433
DB 1609 TTAGCCTGCAAAAGATGCCAATACTAATGGGAGCATGGCTCATAGTCCACTCTCTCTG 1668
QY 434 SerAlaGlnSerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeu 453

Db 1669 TCAGTGCAGTCTGTGATGGGGAGCTGAACACACACCTGTCCAGAGAGTCCACCCCTTG 1728

Qy 454 AlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGlu 473

Db 1729 CCCATCTCTTCGGGAATGACACACGGCTAGAGTGGCTCACTGGCTGAAGTAAAGAG 1788

Qy 474 AsnProProPheTyrGlyValLeuArgTrpIleGlyGlnProProGlyLeuAsnGluVal 493

Db 1789 AACCCCGCTTATCGTGGGTATCCGTTGGATTGGCCAGCCACCGGGCTCAGTGACGTG 1848

Qy 494 LeuAlaGlyLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGly 513

Db 1849 CTAGCTGGACTGGAATGGAATGAATGCGCAGGCTGTACAGATGGAATCTTCAGGGGC 1908

Qy 514 ThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArg 533

Db 1909 AGCGGATTTTACGTGTGCTTGAGAGGACACTGTTGTGAACTGAAGAGCTGCAGA 1968

Qy 534 ProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSer 553

Db 1969 CGGACTCTAGTGTTCATCTTGCAGCCTGTTTCCAAATCAAAATTGAAGGTGTAACCT 2028

Qy 554 LeuAlaPheGlyClyTyrLeuSerGluValValGluGluAsnThrProProlysMetGlu 573

Db 2029 TTAGCATTTGGGGCTATTAAAGTGAAGTAGTAGAAGAAATACCTCCACTAAATGGAA 2088

Qy 574 LysGluGlyLeuGluIleMetIleGlyLysLysLysGlyIleGlnGlyHisTyrAsnSer 593

Db 2089 AAGGAAGTTTAGAGATATGATTGGAAGAGAAGGCAATCAGGGCCATACAACTTCT 2148

Qy 594 CysTyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrVal 613

Db 2149 TGTACTTAGACTCAACTTATTCTGCTATTGCTTTAGTTTCTGCTCCCTGGACACTGTG 2208

Qy 614 LeuLeuArgPro 617

Db 2209 TTACTTAGACCC 2220

RESULT 6

LOCUS BC028885 3137 bp mRNA linear HTC 19-NOV-2003

DEFINITION Mus musculus cylindromatosis (curban tumor syndrome), mRNA (cDNA clone IMAGE:3965911), with apparent retained intron.

ACCESSION BC028885

VERSION BC028885.1 GI:23468224

KEYWORDS HTC.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 3137)

AUTHORS Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altshuler S.F., Ziesler B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Schetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raja S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahy J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A., Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S., Krzywinski M.I., Skalska U., Smalios D.E., Scherch A., Schein J.E., Jones S.J. and Marra M.A.

TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

MEDLINE 22388257

12477932

2 (bases 1 to 3137)

Strausberg,R.

Direct Submission

Submitted (01-MAY-2002) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgabbs-r@mail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Baylor College of Medicine Human Genome

Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: amg@bcm.tmc.edu

Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,

Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,

A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAK Plate: 40 Row: g Column: 11

This clone has the following problem: retained intron.

FEATURES

source

1. 3137

/organism="Mus musculus"

/mol_type="mRNA"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="IMAGE:3965911"

/tissue_type="Mammary tumor. Metallothionien-TGF alpha

model. 10 month old virgin mouse. Taken by biopsy."

/clone_lib="NCI CGAP Mam1"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

ORIGIN

Alignment Scores:

Fred. No.: 1.42e-274 Length: 3137

Score: 2769.00 Matches: 546

Percent Similarity: 66.00% Conservative: 15

Best Local Similarity: 64.24% Mismatches: 16

Query Match: 55.01% Indels: 273

DB: 3 Gaps: 1

US-09-671-687A-3 (1-949) x BC028885 (1-3137)

Qy 373 ValAlaGluAspProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSer 392

Db 186 GTTGCAAGAGACCTCGAAAGTCACCTACAGAGATGCTTCGGACTTCGGACATTCATCT 245

Qy 393 ProProLeuGlnProProValAsnSerLeuThrGluAsnArgPheHisSerLeu 412

Db 246 CCTCCACCGCAGCTCTCTCCATGAACTCTTGTCTAGCGAGAACACAGATTCCTCTTA 305

Qy 413 ProPheSerLeuThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSer 432

Db 306 CCCTTCAGCTGACAAAGATGCCCAATCTAATGGCAGCATGGCTCATAGTCCACTCTCT 365

Qy 433 LeuSerAlaGlnSerValMetGluLeuLeuAsnThrAlaProValGlnGluSerProPro 452

Db 366 CTGTCACTGAGCTGTGTATGGGGAGCTGAACAGCACACTGTTCAGGAGAGTCCACCC 425

Qy 453 LeuAlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLys 472

Db 426 TTGCCCATCTCTTCTGGGAATGCACACGGGCTAGAGGTGGGCTCACTGGCTGAAGTAAA 485

Qy 473 GluAsnProProPheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGlu 492

Db 486 GAGAACCCCGCTCTATGGGTTATCCGTTGGATTGGCCAGCCACCAGGGCTCAGTGAC 545
Qy 493 ValLeuAlaGlyLeuGluLeu----- 499
Db 546 GTGCTAGCTGGACTGGAACCTGTAATTAATTTGGCCAAAAATCTACAACCTCTTTCCCTG 605
Qy 499 ----- 499
Db 606 CATGAGGTTTTGTCTAGATTCCCTCTATAGATAATTTCTAGAAAAATCATCATGTGTG 665
Qy 499 ----- 499
Db 666 TGTGTGTGTGTATGTGCTGTCATGTATGTATTTCCAGTAATAGAGAAAAATATTTCTT 725
Qy 499 ----- 499
Db 726 ACTGTATGGCCTAGCTTTCATTCACCCCTTTGCCANTCCTTAAACTTCTGGCTCATTTTTTC 785
Qy 499 ----- 499
Db 786 CTCGTGTTTCTTGTCTGTATATCAAGGAGTAGAGGAATTTTATAGGCTAGACTATTTAGCT 845
Qy 499 ----- 499
Db 846 TCCTTATATTAAGTATAATGCACCTAACAAAGCAATATCAGTTTTTCTTCTTACTCCTC 905
Qy 499 ----- 499
Db 906 AGCCTCCTCTGATATGATTGGCTTTTACAAATTAATCATCATGTTCAGAGGCTTTTGGTAGA 965
Qy 499 ----- 499
Db 966 AGAGGTTTTCCCTTTGGATATTAGAAAAAGAAATATAGTGTGATATTTCTATAATTTATGAATG 1025
Qy 499 ----- 499
Db 1026 AGGAACCTTGAATAAACTTAGCTTTTACCTTAATGTGCTTTTGAATTAATTTTTTTCATTT 1085
Qy 499 ----- 499
Db 1086 GTTATCTTGATGTGAGGAACCTAGATAAATTAATCTACCTTCCCGCCCTTGATCTTA 1145
Qy 499 ----- 499
Db 1146 ATACTTTTGAAATGGCTCAGGGCAAAATCAGTAAACCAAGAAATGCATGAAGTACTTCT 1205
Qy 499 ----- 499
Db 1206 GGTGGGCTAAAGAAATGTCTTCTTAAACCACTACCTCGTCAGCTAGTCAGCTCCCTTTCT 1265
Qy 499 ----- 499
Db 1266 TGATGCTCTTCTATTTCAGCACAGATTGATGAAGGAATTCAGACAAAGTGTAAAGTTCTT 1325
Qy 500 ----- -G1 500
Db 1326 CCCCACTCCTCTGCAAGGTGCAGTGTAAACACATACCAGAAAGACTTTCTTTTTCAGGA 1385
Qy 500 uAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyrPheThrCysAl 520
Db 1386 AGATGAATCGCAGGCTGTACAGATGGAACCTTTTCAGGGGCACGCGGTATTTTCAGGTGTC 1445
Qy 520 aLeuIysLysAlaLeuPheValLysLeuLysSerCysAtgProAspSerArgPheAlaSe 540
Db 1446 CCTGAAGAAGGCACCTGTTGTGAACCTGAAGAGCTGCAGACCCGACTCTAGGGTTGCAIC 1505
Qy 540 rLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPheGlyGlyTyrLe 560
Db 1506 CTTGCAGCCTGTTTCCAAATCAGATTGAAGGTGTAATCTCTTAGCATTTGGGGGCTATTT 1565
Qy 560 uSerGluValGluGluAsnThrProProLysMetGluLysGluGlyLeuGluLeMe 580
Db 1566 AAGTGAAGTAGTAGAAGAAAAATACTCCACCCTAAATAATGGAAGGAAGGTTTAGAGATAAT 1625

Qy 580 tIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeuAspSerThrIle 600
Db 1626 GATTGGAAAGAAAGGACATCCAGGGCCATTCAAAATTCITGTACTTAGACTCAACATTT 1685
Qy 600 uPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArgProLysGluLyl 620
Db 1686 ATTCTGCTTATTTGCTTTTAGTTCTGCCCTGGACACTGTGTACTTAGACCCCAAGAGNA 1745
Qy 620 sAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGluIleValAsnPr 640
Db 1746 GAATGATATAGAGTATTACAGTGAGACTCAGGAGCTACTGAGGACAGAGATAGTCAATCC 1805
Qy 640 oLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArgLysIleLeuGly 660
Db 1806 TCTGAGAAATATATGGATATGTGTGTCACAAAGATTATGAAGCTGAGGAAAAATACTTGA 1865
Qy 660 uLysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGluGluPheLeuAs 680
Db 1866 AAAAGTTGAGCTGCATCAGGTTTTTACCTCTGAAGAAAAAGATCCTGAAGAAATTTCTTAA 1925
Qy 680 nIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArgSerAlaGlyGly 700
Db 1926 TATCCTGTTTCATGATATTTTAAAGGTTGAACCATTTGTAAAAATAAGATCAGCAGGTCA 1985
Qy 700 nLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGluLysValGlyVa 720
Db 1986 AAAAGTTCAAGACTGTAACTTCTATCAAAATTTTATGMAAAAAAATGAGAAAGTTGGAGT 2045
Qy 720 lProThrIleGlnLeuLeuGluTyrPhePheIleAsnSerAsnLeuLysPheAlaGly 740
Db 2046 ACCCAAAATTCAGCAGTATTAGAAATGTTTATCAACAGCAACCTGAAATTTGCAGA 2105
Qy 740 uAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPheLysLeuPheLyl 760
Db 2106 GGCACCATCATGCTTGATATCCAGATGCTCGGTTTGGAAAAAGACTTTAACTATTATA 2165
Qy 760 sLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAspThrProArgGly 780
Db 2166 AAAAAATTTTCTTCCCTGGAAATTAATATAACAGATTTACTTGAAGACACTCCCGGCA 2225
Qy 780 nCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyrAspAspProAs 800
Db 2226 GTCCCGCATCTGTGGAGGACTTCGATGTACGAGTGCAGGGAGTGCTATGACGATCCGGA 2285
Qy 800 pIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGlnValHisLeuHi 820
Db 2286 CATCTCAGCTGGGAAGATCAAGCAGTTCTGTAAAGCTCTGCAGCACTCAGGTTCACTTCA 2345
Qy 820 sProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAspLeuProAspTyr 840
Db 2346 TCCAGGAGGTTGAATCAATTCATTATCATCCAGTATCACTCCCAAGAGACTTGCCCTGACTG 2405
Qy 840 pAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAlaValLeuCysAl 860
Db 2406 GGACTGGAGACACGGCTGCATCCCTGTCAAGATGGAGTTATTGCTGTCTCTGCTCAT 2465
Qy 860 eGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSerAlaTrpLeuPh 880
Db 2466 AGAAACGAGCCACTATGTGCTTTTGTGAAGTAGTGGGAAGGATGACTCTGCTGGCTTTT 2525
Qy 880 ePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIleProGlnValTh 900
Db 2526 CTTTGACAGCATCGCGGATCGAGATGGTGTGTCAGATGGCTTCAACATTCACCAAGTGAC 2585
Qy 900 rProCysProGluValGlyGlyTyrLeuLysMetSerLeuGluAspLeuHisSerLeuAs 920
Db 2586 GCCTGCCCCAAGAGTGGGAGAGTACTTGAAGATGTCTCTGGAGGACCTGCACTCTTTTGA 2645
Qy 920 pSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyrMetCysMetTy 940
Db 2646 CTCCAGAAGGATTCAGGCTGTGGCGCAGACTTCTTTTGGATGCATACATGTGCATGTA 2705


```

Qy 940 rClnSerProThrMetSerLeuTyrIlys 949
Db 2706 CCAGAGTCCACCAACCATGAGCGCTGACAAA 2733

RESULT 7
BO433523 1063 bp mRNA linear EST 24-MAY-2002
LOCUS AGENCOURT_7766644 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:6015509
DEFINITION 5', mRNA sequence.
ACCESSION BO433523
VERSION BO433523.1 GI:21172599
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1063)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLMI3211 row: 1 column: 06
High quality sequence stop: 672.

FEATURES
source
1..1063
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6015509"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="PH10B (phage-resistant)"
/clone_lib="NIH_MGC_92"
/note="Organ: testis; Vector: pCMV-SPORT6; Site: 1: NotI;
Site: 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN
Alignment Scores:
Pred. No.: 1,97e-151 Length: 1063
Score: 1572.50 Matches: 317
Percent Similarity: 95.27% Conservative: 5
Best Local Similarity: 93.79% Mismatches: 12
Query Match: 31.24% Indels: 7
DB: 5 Gaps: 0

US-09-671-687A-3 (1-949) x BO433523 (1-1063)

Qy 510 ThrPheArgGlyThrArgTyrPheThrCysAlaLeuIlysIysAlaLeuPheValIysLeu 529
Db 1 ACCTTCAGAGGCGACTCGGTATTTTACCTGTGCGCTGAGGAAGCGCTGTTTGTGAACATG 60

Qy 530 LysSerCysArgProAspSerArgPheAlaSerLeuGlnProValSerLysGlnIleGlu 549
Db 61 AAGAGCTGAGGCGCTGACTCTAGTGTGCAATTCAGCGCGGTGTTCCAAATCAGATTGAG 120

Qy 550 ArgCysAsnSerLeuAlaPheGlyGlyTyrLeuSerGluValValGluGluAsnThrPro 569
Db 121 CGCTGTAACTCTTTAGCAITTTGAGGCTACTTAAGTGAAGTACTAGGAAGAAATCTCCA 180

Qy 570 ProLysMetGluLysGluGlyLeuGluIleMetIleGlyLysLysGlyIleGlnGly 589
Db 181 CCAGAAATGAGAAAGAGCTTGGAGATTAATGTTGGGAAGGAAGGATCCAGGGT 240

590 HisTyrAsnSerCysTyrIleuAspSerThrLeuPheCysLeuPheAlaPheSerSerVal 609
241 CATTACAATTTCTGTTTACTTAGACTCAACCTTATTCTGCTTATTGCTTTTAGTTCTGTT 300

610 LeuAspThrValIleuLeuArgProLysGluLysAsnAspValGluTyrTyrSerGluThr 629
301 CTGGACACTGTGTTTACTTAGACCCCAAGAAAGAACGATGTAGATATATTAGTGAACACC 360

630 GlnGluLeuLeuArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCysAla 649
361 CAAGAGCTACTGAGGACAGAAATTTGTAATCTCTCAGATATATGATATGTTGTGTGCC 420

650 ThrLysIleMetLysLeuArgLysIleLeuGluLysValGluAlaAlaSerGlyPheThr 669
421 ACAAAATTTATGAAACTGAGGAAATCTTTGAAAGAGTGAGGCTGCATCAGGATTTTACC 480

670 SerGluGluLysAspProGluGluPheLeuAsnIleLeuPheHisIleLeuArgVal 689
481 TCTGAAGAAAGAAAGATCTGAGGAAATTTTGAATATTCTGTTTTCATCATATTTTAAGGGA 540

690 GluProLeuLeuLysIleArgSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGln 709
541 GAACCTTTGCTAAATAATAGATCAGCAGGTCAAAAGGTACAAAGATTGTTACTTCTATCAA 600

710 IlePheMetGluLysAsnGluLysValGlyValProThrIleGlnGlnLeuLeuGluTrp 729
601 ATTTTATGMAAAATGAGAAAGTTGCGCTTCCACAAATTCAGCAGTGTGTTAGAATGG 660

730 SerPheIleAsnSerAsnLeuLysPheAlaGluAlaProSerCysLeuIleIleGlnMet 749
661 TCTTTTATCAACAGATTAACCTGAAATTTGCAGAGGCCACCATCATGTCTGATTATTCAGATG 720

750 ProArgPheGlyLysAspPheLysLeuPheLysIlePheProSerLeuGluLeuAsn 769
721 CTTGATTTTGGAAAGACTTTAAACTATTTPAAAATTT-TTCTCTCTCTGGG-ATTAAAT 778

770 IleThrAspLeuLeuGluAspThrProArgGlnCysArgIleCysGlyGlyLeuAlaMet 789
779 ATAACAGATTTTACTTGAAGACACTCCACAGACA-TGCCGATATGTGCGAGGCTTGCATATG 837

790 TyrGluCysArgGluCysTyrAspAspPro-AspIleSerAlaGlyLysIleLysGlnPh 809
838 TATGANTGTANAGAANTGCTACGACGATCCGGGACATCTCAGCTGGAATAATCAAGCAGTT 897

809 eCysLysThrCysAsnThrGlnValHisLeuHis-ProLysArgLeuAsnHisLysTyrA 829
898 TTGTAACCTGCAACACTCCAGTCCACCTTTCTCCGAGAGCTGAATCTCTAAATTTT 957

829 snProVal-SerLeuProLysAspLeuProAspTrp-AspTrpArg 843
958 ACCCGTGTGCTTCCCAAGACTACCCGACACTGGGACTGGGAAG 1003

BM457960 1039 bp mRNA linear EST 05-FEB-2002
AGENCOURT_6411653 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5530295
DEFINITION 5', mRNA sequence.
BM457960 1 GI:18507000
VERSION BM457960.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1039)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

```


DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM12210 row: b column: 24
 High quality sequence stop: 702.
 Location/Qualifiers
 source
 1..1039
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5530295"
 /tissue_type="leiomyosarcoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 71"
 /note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 2.1 kb."

ORIGIN

Alignment Scores:
 Pred. No.: 3,44e-151 Length: 1039
 Score: 1570.00 Matches: 312
 Percent Similarity: 91.35% Conservative: 5
 Best Local Similarity: 89.91% Mismatches: 16
 Query Match: 31.19% Indels: 14
 DB: 4 Gaps: 4

US-09-671-687A-3 (1-949) x BM457960 (1-1039)

QY 547 GlnileGluArgCysAsnSerLeuAlaPheGlyGlyTyrLeuSerGluValValGluGlu 566
 |||||
 Db 14 CAGATTGAGCGTGTAACTCTTTAGCATTTGGAGGCTACTTAAGTGAAGTAGTAGAAGAA 73
 |||||
 QY 567 AsnThrProProLysMetGluGlyGluGluLeuMetileGlyLysLysLysGly 586
 |||||
 Db 74 AATACTCCACCAAAATGAAAAAGAGCGCTGAGATATGATGTTGGAGAAAGAAAGGC 133
 |||||
 QY 587 IleGlnGlyHisTyrAsnSerCysTyrLeuAspSerThrLeuPheCysLeuPheAlaPhe 606
 |||||
 Db 134 ATCCAGGGTCATTACATCTTGTTACTTAGACTCAACCTTATCTGCTTATTGCTTTT 193
 |||||
 QY 607 SerSerValLeuAspThrValLeuLeuArgProLysGluLysAsnAspValGluTyrTyr 626
 |||||
 Db 194 AGTCTGTTCTGACACTGTGTTACTTAGACCCAAAGAAAGCAACATGTAGATATTAT 253
 |||||
 QY 627 SerGluThrGlnGluLeuLeuArgThrGluIleValAsnProLeuArgIleTyrGlyTyr 646
 |||||
 Db 254 AGTGAACCCCAAGAGCTACTGAGGACAGAAATTTGTAATCCTCTGAGATATATGGATAT 313
 |||||
 QY 647 ValCysAlaThrLysIleMetLysLeuArgLysIleLeuGluLysValGluAlaAlaSer 666
 |||||
 Db 314 GTGTGTGCCACAAAAATTTAGAACTGAGGAAATTTCTGAAAGGTGGAGGCTGATCA 373
 |||||
 QY 667 GlyPheThrSerGluGluLysAspProGluGluPheLeuAsnileLeuPheHisIle 686
 |||||
 Db 374 GGATTATCTCTGAAGAAAAAGATCTGAGGAATCTTGAATATTCTGTTTCATCATATT 433
 |||||
 QY 687 LeuArgValGluProLeuLeuLysIleArgSerAlaGlyGlnLysValGlnAspCysTyr 706
 |||||
 Db 434 TTAAGGGTGAAGACCTTTGCTTAAATAAAGATACAGCAGGTCAAAAGGTACAAAGTTGTTAC 493
 |||||
 QY 707 PheTyrGlnIlePheMetGluLysAsnGluLysValGlyValProThrIleGlnGlnLeu 726
 |||||
 Db 494 TTCTATCAAAATTTTATGAAAAAATGAGAAATGTTGGCGTTCCCAATTCAGCAGTTG 553
 |||||
 QY 727 LeuGluTyrSerPheIleAsnSerAsnLeuLysPheAlaGluAlaProSerCysLeuIle 746
 |||||
 Db 554 TTAGAATGTCCTTTTATCAACAGTAACTGGAATTTTCAGAGGACCATCATGCTGATT 613
 |||||
 QY 747 IleGlnMetProArgPheGlyLysAspPheLysLeuPheLysLysIlePheProSerLeu 766
 |||||
 Db 614 ATTCAGATGCTCGATTGGAAAAAGACTTTAACTATTTAAAAAATTTTCTCTCTG 773

QY 767 GluLeuAsnIleThrAspLeuLeuGluAspThrProArgGlnCysArgIleCysGlyGly 786
 |||||
 Db 674 GAATTAATAATAACAGATTTTACTTTGAAGACACTCCCGACAGATGCGGATATGTGAGGG 733
 |||||
 QY 787 LeuAlaMetTyrGluCysArgGluCysTyrAspAspProAspIleSerAlaGlyLysIle 806
 |||||
 Db 734 CTTGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 793
 |||||
 QY 807 LysGlnPheCysLysThrCysAsnThrGlnValHisLeuHisProLysArgLeuAsnHis 826
 |||||
 Db 794 AAGCAGTTTTGTAAACCTGCAACACACTCAAGTCCACCTTTCATCCGAGAGGCTGAATCAT 853
 |||||
 QY 827 LysTyrAsn-ProValSerLeuProLysAsp-LeuProAspTyr-AspTyr-ArgHisGly 845
 |||||
 Db 854 AATATAAACCCCAAGTGTCACTTCCCAAGAGACTTTTACCGGACTGGGACTGGNAGACACGG 913
 |||||
 QY 845 YCysIlePro--CysGlnAsnMetGluLeuPheAlaValLeuCysIleGluThrSerHis 864
 |||||
 Db 914 CTGCAATCCCTTGGCCAGCAATATG-----GGAGTTATTTTGTGGTCTCCCGCCA 964
 |||||
 QY 864 sTyrValAlaPheValLysTyrGlyLysAspSerAlaTyrLeuPhePheAspSerMe 884
 |||||
 Db 965 TAGAAAAACAA--GCCACCTATGGT-----TGGCTTTTGGTGAAGTAAAT 1006
 |||||
 QY 884 tAlaAspArgAsp 888
 |||||
 Db 1007 GTGGAAAGGAAC 1019
 |||||
 RESULT 9
 LOCUS AK042764
 DEFINITION Mus musculus 7 days neonate cerebellum cDNA, RIKEN full-length
 enriched library, clone:A730022C13 product:CYLINDROMATOSIS (TURBAN
 TUMOR SYNDROME) homolog [Homo sapiens], full insert sequence.
 ACCESSION AK042764
 VERSION AK042764.1 GI:26335312
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1
 Carninci, P. and Hayashizaki, Y.
 Itoh M., Konno H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 2
 20499374
 11042159
 3
 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 4
 20530913
 11076861
 5
 The RIKEN Genome Exploration Research Group Phase II Team and the
 FANTOM Consortium.
 Functional annotation of a full-length mouse cDNA collection

BI666276	603290787F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:5325140 5',	938 bp	linear	EST 12-SEP-2001
LOCUS	RNA sequence.			
DEFINITION	BI666276			
ACCESSION	BI666276.1	GI:15580509		
VERSION	EST.			
KEYWORDS	Mus musculus (house mouse)			
SOURCE	Mus musculus			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
REFERENCE	NIH-MGC http://mgc.ncl.nih.gov/ .			
AUTHORS	National Institutes of Health, Mammalian Gene Collection (MGC)			
JOURNAL	Unpublished (1999)			
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgabbs@mail.nih.gov Tissue Procurement: Jeffrey Green M.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM1824 row: n column: 21 High quality sequence stop: 776.			
FEATURES	Location/Qualifiers			
source	1..938			
	/organism="Mus musculus"			
	/mol_type="mRNA"			
	/strain="FVB/N"			
	/db_xref="taxon:10090"			
	/clone="IMAGE:5325140"			
	/sex="female, virgin"			
	/tissue type="infiltrating ductal carcinoma"			
	/dev stage="5 months"			
	/lab_host="DH108"			
	/clone_lib="NCI_CGAP Mam6"			
	/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigator providing samples: Jeffrey Green, M.D., NIH"			
ORIGIN				
Alignment Scores:				
Pred. No.:	2,79e-137	Length:	938	
Score:	1435.00	Matches:	277	
Percent Similarity:	93.16%	Conservative:	9	
Best Local Similarity:	90.23%	Mismatches:	18	
Query Match:	28.51%	Indels:	5	
DB:	4	Gaps:	1	
US-09-671-687A-3 (1-949) x BI666276 (1-938)				
QY	637 lleValAsnProLeuArgileTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg	656		
Db	2 ATAGTCAATCTCTGAGAATATATGATATGTGTGCCACAAAGATTATGAACTGAGG	61		
QY	657 LysIleLeuGluLysValGluAlaSerGlyPheThrSerGluLysAspProGlu	676		
Db	62 AAAATCTTGAAGAGTTGAGGCTGCATCAGGTTTACCTCTGAAGAAAGATCTTGAA	121		
QY	677 GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLysIleArg	696		
Db	122 GAATCTTAATATCTCTGTTTCATGATATTTTAAGGTTGAACCATTTTAAATAAGA	181		
QY	697 SerAlaGlyGlnLysValGlnAsp-CysTyrPheTyrGlnIlePheMetGluLysAsnG	716		
Db	182 TCAGCAGGTCAAAAAGTTCAAGACGCTAACTTCTCAAAATTTTATGMAAAAAATGA	241		
QY	716 uLysValGlyValProThrIleGlnLeuLeuGluTrpSerPheIleAsnSerAsnLe	736		
Db	242 GAAAGTTGAGTACCCACAATTACAGAGTTATTAGAAATGCTCTTTTATCAACAGCAAC	301		
QY	736 uLysPheAlaGluAlaProSerCysLeuIleGlnMetProArgPheGlyLysAspPh	756		
Db	302 GAAATTTGCAGAGCACCATCATCTTGATTATCCAGATGCTCGTGTGGAAGACTT	361		
QY	756 eLysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAs	776		
Db	362 TAAACTATTAAAAAATTTTCTCTCCCTGGAATTAATATACAGATTACTTTGAAGA	421		
QY	776 pThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTy	796		
Db	422 CACTCCAGGAGTGGCGCATCTCTGGAGGACTTTCGATGTACGAGTGCAGGAGTGCTA	481		
QY	796 rAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGl	816		
Db	482 TGACGATCCGGACATCTCAGCTGGGAAGATCAAGCAGTTCTGTAAAGACTCGACACTCA	541		
QY	816 nValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAs	836		
Db	542 GGTTCACCTTCATCCAGGAGTTGAATCATCTTATCATCCAGTATCATTCCCAAGA	601		
QY	836 pLeuProAspTrpAspTrpArgHisGlyCysIlePheProCysGlnAsnMetGluLeuPheAl	856		
Db	602 CTTGCTGACTGGGACTGGAGACACGCGCTGCATCCCTGTGAGAAGATGGAGTTATTTC	661		
QY	856 aValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyr-GlyLysAspAspS	876		
Db	662 TGTTCCTGTCATGAAACGAGCCACTATGTAGCTTTTGTGAAGTATGGGGAAGGATGACT	721		
QY	876 erAlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnI	896		
Db	722 CTGCTGGCTTTTCTTTGACAGATGGCGGATCAGATGGTGGTCAAGATGGTTCAAGA	781		
QY	896 leProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspL	916		
Db	782 TTCACAGATGACGCCCTGCCAGAGTGGGAGAGTACTGAAGATGCTCTCGGGAAGAAC	841		
QY	916 eutHisSerLeuAspSerArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlar	936		
Db	842 TGACACCTTGGGACCC-AAAAGAATCCAGGCGC---GGCCCCAACA-CTAGGCGAGGGAC	896		
QY	936 yrMetCysMetTyrGln 941			
Db	897 ACATGTGCTGTACCAG 913			
RESULT 11				
AK048183	2218 bp	linear	HTC 03-APR-2004	
LOCUS	Mus musculus 16 days embryo head cDNA, RIKEN full-length enriched			
DEFINITION	library, clone:cl3009D01 product:CYLINDROMATOSIS (TURBAN TUMOR SYNDROME) homolog (Homo sapiens), full insert sequence.			
ACCESSION	AK048183			
VERSION	AK048183.1	GI:26092705		
KEYWORDS	HTC; CAP trapper.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Mus musculus			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
AUTHORS	Carninci, P. and Hayashizaki, Y.			
TITLE	High-efficiency full-length cDNA cloning			
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)			
MEDLINE	99279253			
PUBMED	10349636			
REFERENCE	2			
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komoto, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.			
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes			
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)			
MEDLINE	20499374			
PUBMED	11042159			
REFERENCE	3			

AUTHORS	Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kusunagi,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
	RIKEN Integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multichannel sequencer
	Genome Res. 10 (11), 1757-1771 (2000)
	20530913
JOURNAL MEDLINE	11076861
	PUBMED
REFERENCE	4
	THE RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
JOURNAL	Functional annotation of a full-length mouse cDNA collection
	Nature 409, 685-690 (2001)
REFERENCE	5
	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
JOURNAL	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
	Nature 420, 563-573 (2002)
REFERENCE	6 (bases 1 to 2218)
	Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
JOURNAL	Direct Submission
	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsr.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
	Please visit our web site for further details. URL:http://genome.gsc.riken.jp/ URL:http://fantom.gsc.riken.jp/
FEATURES	Location/Qualifiers
	1..2218
source	/organism="Mus musculus"
	/mol_type="mRNA"
	/strain="C57BL/6J"
	/db_xref="FANTOM,DB:C130039D01"
	/db_xref="taxon:10090"
	/clone="C130039D01"
	/tissue_type="head"
	/clone_lib="RIKEN full-length enriched mouse cDNA library"
	/dev_stage="16 days embryo"
	1..2218
	/note="CYLINDROMATOSIS (TURBAN TUMOR SYNDROME) homolog [Homo sapiens] (SPTP Q96EH0, evidence: FASTA, 93.8%ID, 30.5%length, match=873)"
misc_feature	1..2218
	/note="CYLINDROMATOSIS (TURBAN TUMOR SYNDROME) homolog [Homo sapiens] (SPTP Q96EH0, evidence: FASTA, 93.8%ID, 30.5%length, match=873)"
	Alignment Scores: 1.03e-135 Length: 2218
	Pred. No.: 1426.00 Matches: 273
ORIGIN	Score: 96.56% Conservative: 8
	Percent Similarity:
Best Local Similarity:	93.81% Mismatches: 10
	Query Match: 28.33% Indels: 0
DB:	Gaps: 0
	US-09-671-687a-3 (1-949) x AK048183 (1-2218)
QY	406 GluA5A9A9PheHisSerLeuProPheSerLeuThrLysMetProAenThrAenGlySer 425
	1 AAGAACAGATATCCACCTCTTACCTTCCAGCTGACAAAGATGCCCAATCTAATATGGCAGC 60
QY	426 IleGlyHisSerProLeuSerLeuAlaGlnSerValMetGluGluLeuAenThrAla 445
	61 ATGGCTCATAGTCCACTCTCTCTGTGATGGGGAGCTGAAACGACACA 120
QY	446 ProValGlnGluSerProProLeuAlaMetProGlyAsnSerHisGlyLeuGluVal 465
	121 CTGTCCAGGAGATCCACCTTGGCCATCTCTTCTGGGAATGCACACGGGCTAGAGGTG 180
QY	466 GlySerLeuAlaGluValLysGluAenProPheTyrGlyValIleAgtPrlleGly 485
	181 GGCTCACTGGCTGAAGTAAAGAGAAACCCCGCTTCTATGGGGTTATCCGTGGATTGGC 240
QY	486 GlnProGlyLeuAenGluValLeuAlaGlyLeuGluLeuGluAenGluCysAlaGly 505
	241 CAGCCACCAAGGGCTCAGTGACGTGTGGACTGGAACTGGAAGATGAATGAGCAGGC 300
QY	506 CysThrAspGlyThrPheArgGlyThrArgTyrPheThrCysAlaLeuLysLysAlaLeu 525
	301 TGTACAGATGGAACTTTCAGGGGCAAGCGGTATTTACGTGTGCCCTGAGAGGCACTG 360
QY	526 PheValLysLeuLysSerCysArgProAspSerArgPheAlaSerLeuGlnProValSer 545
	361 TTTGTGAAACTGAAGAGCTGCACAGCGGACTCTAGGTTCATCTTGCACGCTGTTTCC 420
QY	546 AenGlnIleGluArgCysAenSerLeuAlaPheGlyGlyTyrLeuSerGluValValGlu 565
	421 AATCAAAATGAAAGGTGTAACCTCTTTAGCATTTTGGGGCTATTAACTGAAGTAGTAGAA 480
QY	566 GluA5A9A9PheHisSerLeuProPheSerLeuThrLysMetIleGlyLysLysLys 585
	481 GAAATATCTCCACTTAATGGAAGAGAGTTTGAAGATATATGATGGAAGAGAGAGAA 540
QY	586 GlyIleGlnGlyHisTyrAenSerCysTyrLeuA5A9A9PheThrLeuPheCysLeuPheAla 605
	541 GGCAATCCAGGCGCATTAACAATCTTGTACTTAGACTCAACTTTATTCCTATTGCT 600
QY	606 PheSerSerValLeuA5A9PheThrValLeuLeuArgProLysGlyLysAenAspValGluTyr 625
	601 TTTAGTTCTGCGCTGGACACTGTGTACTTAGACCCCAAGAGAGATGATATAGAGTAT 660
QY	626 TyrSerGluThrGlnGluLeuLeuArgThrGluIleValAenProLeuArgIleTyrGly 645
	661 TACAGTGAGACTCAGGAGCTTACTGAGACAGAGATGATCACTCTCTGAGAAATATATGGA 720
QY	646 TyrValCysAlaThrLysIleMetLysLeuArgLysIleLeuGluLysValGluAlaAla 665
	721 TATGTGTGTGCCACAAAGATTATGAAACTGAGGAGAAATACTTGAAAAAGTTGAGCGTGCA 780
QY	666 SerGlyPheThrSerGluGluLysAspProGluGluPheLeuAenIleLeuPheHisHis 685
	781 TCAGGTTTACCTCTCAGAGAGAAAGATCTCTGAGAGAAATCTTAAATATATCTCTTTCATGAT 840
QY	686 IleLeuArgValGluProLeuLeuLysIleArg 696
	841 ATTTTAAAGGTTGAACCAATGTTTAAAAAATAAGG 873
RESULT 12	BM480127 1133 bp mRNA linear EST 05-FEB-2002
	BM480127
LOCUS	AGENCOURT_6424196 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5491647
	5', mRNA sequence.
DEFINITION	BM480127
	ACCESSION
VERSION	BM480127.1 GI:18529169

Fax: 319 335 9565

Email: bento-soares@uiowa.edu

Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).

Seq primer: M13 REVERSE.

FEATURES

source

Location/Qualifiers
1. 812
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-FJO-qga-C-11-0-UI"
/tissue_type="embryo"
/dev_stage="embryo"
/lab_host="PH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-R-FJO"
/note="Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I;
UI-R-FJO is a cDNA library containing the following
tissue(s): rat embryo. The library was constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:1791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pTVT3-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is CATCTCTACT. This library
was created for the University of Iowa Program for Rat
Gene Discovery and Mapping (Val Sheffield, Bento Soares
and Tom Casavant)"

ORIGIN

Alignment Scores:

Pred. No.: 4,19e-132 Length: 812
Score: 1384.00 Matches: 256
Percent Similarity: 98.48% Conservative: 4
Best Local Similarity: 96.97% Mismatches: 4
Query Match: 27.49% Indels: 0
DB: 6 Gaps: 0

US-09-671-687A-3 (1-949) x CAS12526 (1-812)

627 SerGluThrGlnGluLeuLeuArgThrGluLeuValAsnProLeuArgIleTyrGlyTyr 646
:::|||||
20 ACGGAGACTCAAGAGCTACTGAGGACAGAGATAGTCAATCTCTGAGAATATATGGATAT 79
647 ValCysAlaThrIysIleMetIysLeuArgIleIleIeuGluIysValGluAlaIleSer 666
:::|||||
80 GTGTGTGCCACAAAGATTATGAGCTGAGGAAATACTTGAAAAGTTGAGCGTGCATCA 139
667 GlyPheThrSerGluGluIysAspProGluGluPheLeuAsnIleLeuPheHisIle 686
:::|||||
140 GAATTTACTCTGAGGAAAAGATCTCTGAGAATTTCTAAACATCTCTTTTCATGATATT 199
687 LeuArgValGluProLeuLeuIysIleArgSerAlaGlyGlnIysValGlnAspCysTyr 706
:::|||||
200 TTAAGGTTTGAACCATTTTAAATAAAGTCAAGCTCAGAGGTCAGAAAGTTCAAGACTGTAA 259
707 PheTyrGlnIlePheMetGluIysAsnGluIysValGlyValProThrIleGlnGlnLeu 726
260 TTCTATCAAAATTTTATGGAATAAATGAGAAAGTCGGAGTACCCACAATCCAGCATTA 319
727 LeuGluTyrSerPheIleAsnSerAsnLeuIysPheAlaGluAlaProSerCysLeuIle 746
320 TTAGAATGGTCTTTATCAACACCACTGAAATTTGCAGAGGACCAATCATCTGCTGATT 379
747 IleGlnMetProArgPheGlyIysAspPheIysLeuPheIysLysIlePheProSerLeu 766
:::|||||

Db 380 ATCCAGATGCCTCGGTTTGGGAAGACTTTAAACTATTAAAAAATTTTCTCTCCTG 439
Qy 767 GluLeuAsnIleThrAspLeuLeuGluAspThrProArgGlnCysArgIleCysGlyGly 786
Db 440 GAATTAATATAACAGATTACTTTGAAGACACTCCAGGCGAGTCCGCACTCTGTGGAGGA 499
Qy 787 LeuAlaMetTyrGluCysArgGluCysTyrAspAspProAspIleSerAlaGlyIle 806
Db 500 CTCGCCATGTATGAGTGTAGAGAGTCTATGATGACCCGACATCTCGGAGGGAAGATC 559
Qy 807 LysGlnPheCysLysThrCysAsnThrGlnValHisLeuHisProLysArgLeuAsnHis 826
Db 560 AAGCAGTTCTGTAAAGACCTGCAGCACTCAGTTCACTTCAATCCAGAAAGACTGAATCAC 619
Qy 827 LysTyrAsnProValSerLeuProLysAspLeuProAspTrpAspTrpArgHisGlyCys 846
Db 620 ACTTACCATCCAGTATCACTTCCCAAGACTTCCCGAGCTGGGACTGGAGACACGGCTGC 679
Qy 847 IleProCysGlnAsnMetGluLeuPheAlaValLeuCysIleGluThrSerHisTyrVal 866
Db 680 ATCCCGTGTGAGAAGATGGAGTTATTGCTGTCTCTGCATAGAAACCGCCACTATGTT 739
Qy 867 AlaPheValIysTyrGlyLysAspAspSerAlaTyrLeuPheAspSerMetAlaAsp 886
Db 740 GCTTTTGTGAAGTACGGAAGGATGACTCTGCCTGGCTCTTTTGACAGATGGCTGAT 799
Qy 887 ArgAspGlyGly 890
Db 800 CGAGATGGTGGT 811
RESULT 14
BC024596
LOCUS
DEFINITION
Mus musculus cyclinomatosis (turban tumor syndrome), mRNA (cdna
clone IMAGE:4022147), with apparent retained intron.
ACCESSION
BC024596
VERSION
BC024596.1
KEYWORDS
GI:19353300
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
REFERENCE
1 (bases 1 to 1892)
AUTHORS
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.D.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullaby, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalhus, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
JOURNAL
MEDLINE
PUBMED
12477932
REFERENCE
2 (bases 1 to 1892)
AUTHORS
Strausberg, R.
Direct Submission
Submitted (01-MAR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulsegue, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 41 Row: e Column: 18
 This clone has the following problem: retained intron.

FEATURES
source

Location/Qualifiers
 1..1892
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="Czech II"
 /db_xref="taxon:10090"
 /clone="IMAGE:4022147"
 /tissue_type="Mammary tumor metastasized to lung."
 MMTV-LTR/Wnt1 model. Expression driven by an MMTV-LTR enhancer."
 /clone_lib="NCI CGAP_Lu30"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"

ORIGIN

Alignment Scores:
 Pred. No.: 2,65e-129 Length: 1892
 Score: 1363.00 Matches: 246
 Percent Similarity: 98.81% Conservative: 3
 Best Local Similarity: 97.62% Mismatches: 3
 Query Match: 27,08% Indels: 0
 DB: 3 Gaps: 0

US-09-671-687A-3 (1-949) x BC024596 (1-1892)

QY 698 AlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGluLys 717
 DB 1 GCAGGTCAAAAAGTTCACAGCTGTACTTCTATCAAAATTTTATGGAATAAATGAGAAA 60
 QY 718 ValGlyValProThrIleGlnLeuLeuGluTyrSerPheIleAsnSerAsnLeuLys 737
 DB 61 GTTGAGGTACCCACAAATTCAGCAGTATTAGAAATGCTTTTATCAACAGCAACCTGAAA 120
 QY 738 PheAlaGluAlaProSerCysLeuIleGlnMetProArgPheGlyLysAspPheLys 757
 DB 121 TTTCAGAGGACCAATCATGCTGTGATTATCCAGATGCCCTCGGTTTGAAAAGACTTTAAA 180
 QY 758 LeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAspThr 777
 DB 181 CTATTTAAAAAATTTTCTCCCTCGGAATTAATATACAGATTACTTGAAGACACT 240
 QY 778 ProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyrAsp 797
 DB 241 CCCAGGCAGTCGCATCTGTGGAGGACTTGGATGTACGAGTGCAGGAGGTCTATGAC 300
 QY 798 AspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGlnVal 817
 DB 301 GATCCGGACATCTCAGCTGGGAAGATCAAGCAGTTCTGTAAAGACCTGCAGACTCAGGTT 360
 QY 818 HisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAspLeu 837
 DB 361 CACCTTCATCCCGAGGAGGTGAATCTTCTATCATCCAGTATCACTTCCCAAGACTTG 420
 QY 838 ProAspTTPAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAlaVal 857

Db 421 CTTGACTGGAGCTGGAGACACCGCTGCATCCCTGTGCAAGATGGAGTTATTGCTGTT 480
 QY 858 LeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSerAla 877
 Db 481 CTCTGCATAGAAACGACCACTATGTTGCTTTTGTGAAGTATGGGAAGGATGACTCTGCC 540
 QY 878 TrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIlePro 897
 Db 541 TGGCTTTTCTTTTACACAGCATGGCGGATCGAGATGGTGGTCTGAGATGGCTTCAACATTTCA 600
 QY 898 GlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLeuHis 917
 Db 601 CAAGTACGCCCTGCCAGAGTGGGAGAGTACTTTGAAGATGCTCTGGAGAGCTGCAC 660
 QY 918 SerLeuAspSerArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyrMet 937
 Db 661 TCTTTGATTCACAGAGGATTCAGGCTGTGGCGCAGACTTCTTTGGATGCATACATG 720
 QY 938 CysMetTyrGlnSerProThrMetSerLeuTyrLys 949
 Db 721 TGCATGTACCAGATCCCAACCATGAGCTGTACAAA 756
 RESULT 15
 BG867631
 LOCUS
 DEFINITION
 mRNA sequence.
 ACCESSION
 BG867631
 VERSION
 BG867631.1 GI:14218171
 KEYWORDS
 EST.
 SOURCE
 Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE
 1 (bases 1 to 837)
 AUTHORS
 NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL
 Unpublished (1999)
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: LLAM10818 row: 1 column: 03
 High quality sequence stop: 756.
 FEATURES
 Location/Qualifiers
 1..837
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:4913426"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NCI CGAP_SG2"
 /note="Organ: salivary gland; Vector: pCMV-SPORT6; Site 1: NCI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.3 kb. Constructed by Life Technologies. Note: this is a NCI CGAP Library."

ORIGIN

Alignment Scores:
 Pred. No.: 1,59e-127 Length: 837
 Score: 1340.00 Matches: 254
 Percent Similarity: 95.26% Conservative: 7
 Best Local Similarity: 92.70% Mismatches: 12
 Query Match: 26,62% Indels: 2
 DB: 4 Gaps: 0

US-09-671-687A-3 (1-949) x BG867631 (1-837)

```
Qy 611 AspThrValLeuLeuArgProLysGluLysAsnAspValGluTyrTyrSerGluThrGln 630
Db 2 GACACTGTGTTACTTAGACCCAAAGAGAAGATGATATAGAGTATTACAGTGAGACTCAG 61
Qy 631 GluLeuLeuArgThrGluLeuValAsnProLeuArgLleTyrGlyTyrValCysAlaThr 650
Db 62 GAGCTACTGAGGACAGAGATAGTCAATCTCTGAGATATATGAGATATGTGTGCCACA 121
Qy 651 LysIleMetLysLeuArgLysIleLeuGluLysValGluAlaAlaSerGlyPheThrSer 670
Db 122 AAGATTATGAAGAACTGAGGAATACTTGAAAAGTTGAGGCTGCATCAGGTTTACCTCT 181
Qy 671 GluGluLysAspProGluGluPheLeuAsnIleLeuPheHisIleLeuArgValGlu 690
Db 182 GAAGAAAAGATCCTGAAGATTCTTAATATCCTGTTTCATGATATTTTAAGGGTTGAA 241
Qy 691 ProLeuLeuLysIleArgSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIle 710
Db 242 CCATTGTTAAAAATAAGATCAGCAGGTCAAAAAGTTCAAGACTGTAACCTCTATCAAAAT 301
Qy 711 PheMetGluLysAsnGluLysValGlyValProThrIleGlnGlnLeuLeuTrpSer 730
Db 302 TTTATGGAATAAATGAGAAAGTTGAGTACCCCAATTCAGCAGTTATTAGAATGGTCT 361
Qy 731 PheIleAsnSerAsnLeuLysPheAlaGluAlaProSerCysLeuIleIleGlnMetPro 750
Db 362 TTTATCAACAGCAACTGAAATTTGAGAGGGGACCACATCATGTTGATATCCAGATGCCT 421
Qy 751 ArgPheGlyLysAspPheLysLeuPheLysIleIlePheProSerLeuGluLeuAsnIle 770
Db 422 CGGTTTCGGAAGAAAGACTTTAAACTATTAAAAAATT-TTTCCTTCCCTGGAAATTAATATA 480
Qy 771 ThrAspLeuLeuGluAspThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyr 790
Db 481 ACAGATTTTACTTGAAGACACTCCAGGCAGTCCGCATCTGTGGAGGACTTGCATGTAC 540
Qy 791 GluCysArgGluCysTyrAspAspProAspIleSerAlaGlyLysIleLysGlnPheCys 810
Db 541 GAGTGCAGGGAGTGCTATACCATCCGGACATCTCAGCTGGGAAGATCAAGCAGTTCTGT 600
Qy 811 LysThrCysAsnThrGlnValHisLeuHisProLysArgLeuAsnHisLysTyrAsnPro 830
Db 601 AAGACCTGCAGCACTCAGGTTCACTTCACTCCAGGAGGTTGAATCATCTTATCATCCA 660
Qy 831 ValSerLeuProLysAspLeuProAspTrpAspTrpArgHisGlyCysIleProCysGln 850
Db 661 GTATCACTTCCCAAGACTTGCTGACTGGGACTGGAGACACGGCTGCATCCCTGTGAGA 720
Qy 851 AsnMetGluLeuPheAlaValLeuCysIleGluThrSerHisTyrValAlaPheValLys 870
Db 721 AGATGGAAGTTATTGCTGTTCTCTCATAGAAACGAGGCACCTATGTCGCTTTGTGAAG 780
Qy 871 TyrGlyLys-AspAspSerAlaTrpLeuPhePheAspSer 883
Db 781 TATGGGAACGGATGACTCTGGCTGGCTTTCTTTCGACAG 820
```

Search completed: April 18, 2005, 20:07:09
Job time : 6231 secs

345 DNBRLVOVSGDEEOVPSAPKPA
PBBRTTPVSRPOOPETEORN-----SRSMKPSF 393

[illegible]

Db	7	KSISLIFCHSGSSVAESN---SRUYGPNQMHMPKGGVSGALYDNRRLVQYSGDBEQ	64
QY	359	-----QPOSKSNTWIDEVAEDPAKSLTEISTDFDR--SSPPLQPP	398
Db	65	YRSAPKAPRERIIPVSRQOPEIEQRN-----SRSMKPSBDYNTYSTHPPRP	113
QY	399	PVNSLT--TENRPHSLPFLSTKMPNT---NGSIGHSPLSLSAOSVMEELNTAPVQSSPPL	453
Db	114	PSSMNPYMSNTHSLQPSRSKSVQTIQRNQFVRQPAERNERNVESQLNFRIGDQC--	170
QY	454	AMPNGSHGLE---VGSIAEVKENPFY-----GVTRWLGQPPGLNEVL	494
Db	171	-----IWNNGAERGIIKYGFLKGHKTLVAGSHVVVHGLCAQRTGVVRNLGE	219
QY	495	AGLEDEDEACAGCTDGTFRGTRYFTCALXKALFKVLKSCRPD-----SRFAS	540
Db	220	---EIDDE-----GDR-----IVRSAI VOLDIEDVPTAWRRSHDASSYTGSPVVG	261
QY	541	-LQPVSNQIERCNSLAFGGYLSVVEENTPPK-----MEKE-----	575
Db	262	VLVPIQSLKHNTHSGSSNG---ASVQTSITYPTSTQYNNISRRRTDFGMSDGSVEKQKCI	319
QY	576	--GLEMTIGKKKGIQGHYNSCYLDSLTFCLFAPSSFVLDTVLLRPKEKNDVEYVSETOELL	633
Db	320	AKDMQOLVGRQKGIQGVCHNSCYLDATLYAMFVQTTCTDFLL--EKSIKSETAQOQKIL	377
QY	634	RTEIVNPLRIYGVCAIKMKLRKIL-EKVEAASGFTSEKDPPEEFNLILFHHILRVEPL	692
Db	378	AHEIVFPLRKVHYRADHYMKRLKLLAELMHPVHTGLTNEEKDPEEILGFIISKVFHAEPF	437
QY	693	LKIRSAGQVQDCYFYQIEMEKNEKVGVPITIQOLLEWSFNSNLKPAEAPSCLLIOMPRF	752
Db	438	IKLIGONHAKDSYLVPIVVD-DWLGAATSQHLHERMSAQVTFKAKAPVLIMQLPRY	496
QY	753	GKDFLKFKIIFPSLELNITDLLBETPRQRCICGGLAMYECRECVDDPDISAGKIKQFCKT	812
Db	497	GQ-Q-KVFDKILPLETIDITPFVAGAVPACSKQACSEVCPTCFLTRRVFVSEV-IFCRK	554
QY	813	CNTQVHLHKPLRNHKYN---PVSLPKDLPDWDWRHGCIPQNMELFAVLCIETSHYVAFV	869
Db	555	CFHHTLLPEIEDHKSRDLYPPGKPKKPH-----SHKMVLSAVLCIETSHYVAVV	605
QY	870	KYGGKDSAWLFFDSMADRDGGQGNFIPOVTPCPEVGEVWKMS-----LEDLH	917
Db	606	R--TSSNQWVFFDSMADREGLSDGFNVVPVRECGNMSDWLSLQGNRLKDADECQIKVM	663
QY	918	-----SLDSRRIQ-----GCARLLCDAVCMYQ--SPTMS	946
Db	664	ELMKNISPNDRKLEIAMFGQSSLDPLVGLRLSDSYICFYEDASPTSS	711
RESULT 3			
A43336			
microtubule-vesicle linker CLIP-170 - human			
C;Species: Homo sapiens (man)			
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 05-Nov-1999			
R;Accession: A43336			
R;fierre, P.; Scheel, J.; Richard, J.E.; Kreis, T.E.			
Cell 70, 887-900, 1992			
A;Title: CLIP-170 links endocytic vesicles to microtubules.			
A;Reference number: A43336; MUID:92405160; PMID:1356075			
A;Accession: A43336			
A;Status: preliminary			
A;Molecule type: mRNA			
A;Residues: 1-1392 cPIE>			
A;Cross-references: GB:M97501; NID:g180621; PIDN:AAA35693.1; PID:g180622			

```
Db      85  GQWAGIVLDEPIGKNDGSAVGR-YFQCEP-----LKGIF-----TPSKL---TRK 127
QY      320  VDGKSSSHNKPATGSDPGNRRSELFTYLGSSVDSPQSKSNWTWYIDVADPAK 379
Db      128  V-----QAEDENGLQTTPASRATSPCTSTASVSSPSFSPN-----IPQKPSQA- 175
QY      380  SLTEISTDFDRSSPPLQPPVNSLTITENRPHSLPFLSKTQMPNTNGSIGHSPLSLSAQSV 439
Db      176  -----AKPSATPPISNLT-----KTASESIS 197
QY      440  EELNTAPVOESPPLAMPNGNSHGLEVSLAEVKENPFYGVIRWIGOPGLNEVLAGLEL 499
Db      198  NLSEAGSIKK-----GERELKIGDRVLVGGTKA--GVRFLETDFAKGEWCGVEL 246
QY      500  EDEACAGTGTGTRGTRYFTCALKKALFVKLKSRCRPSRFPASLQPVSNQIERCNSLAFGGY 559
Db      247  -DEPLGKNDGAVAGTRYFOCPKGLFA-----PVHKVTKIGF--- 283
QY      560  LSEVVEENTPPKMEKGL-EIMIGKKKGIQCHYNSCYLDSTLFCLEAFSSVLDTVLLRPK 618
Db      284  -----PSTTPAKAKANAVRRVMTTASLKRSPSASSLS-----MSSVASSVSRP- 330
QY      619  EKNDVEYSETQELLRTIENVPLRIYGVYCATKIMK-----LRKILEKVEAASGP 668
Db      331  -----SRTGLTSTSSRYARKISGTTALQEALEKKEQKHIEQLLAERDLERAFAKA- 381
QY      669  TSEKDPPEEFLNIL-----FHHIL-----RVPLLKIRSAQGVQDC 705
Db      382  TSHVGEIEQELALARDHQHVLLEAKMDQLRTWVEADREKVELLNQLNEEKKVEDL 441
QY      706  YF 707
Db      442  QF 443

RESULT 4
restin - human
C:Species: Homo sapiens (man)
C>Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 09-Jul-2004
C:Accession: S22695; S19853
R:Biode, G.; Delabate, J.; Bruengen, J.; Richener, H.; Asselbergs, F.A.M.; Cerletti, N.;
EMBO J. 11, 2103-2113, 1992
A:Title: Restin: a novel intermediate filament-associated protein highly expressed in th
A:Reference number: S22695; MUID:92289675; PMID:1600942
A:Accession: S22695
A:Molecule type: mRNA
A:Residues: 1-1427 <BIL>
A:Cross-references: UNIPROT:P30622; EMBL:X64838; NID:g35998; PID:CAAA6050.1; PID:g35999
C:Keywords: cytoskeleton

Query Match      3.0%; Score 151; DB 2; Length 1427;
Best Local Similarity 20.5%; Pred. No. 0.089;
Matches 99; Conservative 55; Mismatches 171; Indels 157; Gaps 20;

QY      260  GYFVGVDMDNPTGNWDGDFGLVCAPVESTILLHINDIIPSVTQERPPKLAFMERG 319
Db      85  GQWAGIVLDEPIGKNDGSAVGR-YFQCEP-----LKGIF-----TPSKL---TRK 127
QY      320  VDGKSSSHNKPATGSDPGNRRSELFTYLGSSVDSPQSKSNWTWYIDVADPAK 379
Db      128  V-----QAEDENGLQTTPASRATSPCTSTASVSSPSFSPN-----IPQKPSQA- 175
QY      380  SLTEISTDFDRSSPPLQPPVNSLTITENRPHSLPFLSKTQMPNTNGSIGHSPLSLSAQSV 439
Db      176  -----AKPSATPPISNLT-----KTASESIS 197
QY      440  EELNTAPVOESPPLAMPNGNSHGLEVSLAEVKENPFYGVIRWIGOPGLNEVLAGLEL 499
Db      198  NLSEAGSIKK-----GERELKIGDRVLVGGTKA--GVRFLETDFAKGEWCGVEL 246
QY      500  EDEACAGTGTGTRGTRYFTCALKKALFVKLKSRCRPSRFPASLQPVSNQIERCNSLAFGGY 559
```

```
Db      247  -DEPLGKNDGAVAGTRYFOCPKGLFA-----PVHKVTKIGF--- 283
QY      560  LSEVVEENTPPKMEKGL-EIMIGKKKGIQCHYNSCYLDSTLFCLEAFSSVLDTVLLRPK 618
Db      284  -----PSTTPAKAKANAVRRVMTTASLKRSPSASSLS-----MSSVASSVSRP- 330
QY      619  EKNDVEYSETQELLRTIENVPLRIYGVYCATKIMK-----LRKILEKVEAASGP 668
Db      331  -----SRTGLTSTSSRYARKISGTTALQEALEKKEQKHIEQLLAERDLERAFAKA- 381
QY      669  TSEKDPPEEFLNIL-----FHHIL-----RVPLLKIRSAQGVQDC 705
Db      382  TSHVGEIEQELALARDHQHVLLEAKMDQLRTWVEADREKVELLNQLNEEKKVEDL 441
QY      706  YF 707
Db      442  QF 443

RESULT 5
T42734
cytoplasmic linker protein CLIP-115 - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C:Accession: T42734
R:de Zeeuw, C.I.; Hoogenraad, C.C.; Goedknegt, E.; Hertzberg, E.; Neubaer, A.; Grosvel
Neuron 19, 1187-1199, 1997
A:Title: CLIP-115, a novel brain specific cytoplasmic linker protein, mediates the local
A:Reference number: 222252; MUID:98087115; PMID:9427243
A:Accession: T42734
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1046 <DEZ>
A:Cross-references: UNIPROT:O55156; EMBL:AJ000485; NID:g2792008; PIDN:CAA04123.1; PID:g2
C:Genetics: CLIP-115
C:Function:
A:Description: supposed to operate in the control of brain-specific organelle translocat
A>Note: enriched in the dendritic lamellar body (DLB)

Query Match      3.0%; Score 149.5; DB 2; Length 1046;
Best Local Similarity 20.9%; Pred. No. 0.069;
Matches 115; Conservative 66; Mismatches 189; Indels 179; Gaps 24;

QY      309  RPKLAFMMSGVGDGKSSSHNKPATGSDPGNRRSELFTYLGSSVDSPQ 361
Db      8  KPP-----GRGKHSPPVGRPSI-----GSASSVVASASGSKGSLHKQASGPS 53
QY      362  SKSKNTWYIDVADPAKSLTEISTDF-----DRSSPPLQPPPVNSLTITENRF----- 409
Db      54  SAGATT-----TVSEKPGKAAEVGDDFLGDFVVGGERVWVNGVKPCVVQYLGETQFAPQW 109
QY      410  -----HSLPFSITKMP--NTNGSIGHSPLSLSAQS 437
Db      110  AGVVLDDPVGKNDGAVGGLRYFECPALQGIFTRPSSKLTRQPAAGSGSDGHSVESLTAQN 169
QY      438  VMEELNTAPVOESPPL-----AMPNGNSHGLEVSLAEVKENPFYF 478
Db      170  LSLHSGTA-----TPPLTGRVPIRESVNSVKTGNSGSLSDSGSVKRGDKDLHLGDR 225
QY      479  -----GVIRWIGOPGLNEVLAGLEDEACAGTGTGTRGTRYFTCALKKALFVKL- 529
Db      226  VLVGGTKTVRVYVGTDFPAKGEWCGVEL-DEPLGKNDGAVAGTRYFOCPKGLFAFIH 284
QY      530  -----KSCRPDSRPA-----SLQPVSNQIERCNSLA--FGYILSEVVEENTP 569
Db      285  KVIRIGFPSTSPAKAKTKMAMGVSAALTHSPSSSISSVSSVASSVSGRPS----- 336
QY      570  PKMEKGLIEMIGKKKGIQCHYNSCYLDSTLFCLEAFSSVLDTVLLRP--KEKND-VEYY 626
Db      337  -----RSGLLTSTSRVA-----RKISGTTALQEALEKKEQKHIEQL 372
```


Db 1658 SYNEQSGIFAGGQGYRIRVNGKTGLVGAUVSD-ADSKNLLKTSIWHKDIQNHASAAA 1716
QY 380 SUTESTDTRSPPIOPPPVNSLTTFNPHSLTKMPNTNGSIGHSPLSL-----433
Db 1717 SALGLSGGFSYSPKP-----TSGQYS-----TKAEABIGKIGKGPVSLMRFDQV 1760
QY 434 SAQSVMEELNAPVQBS-----PPLAMPPGNHGLEVG-----466
Db 1761 SAKD--DELNEKYRSRIEKGTFKEANLNQNNAGGLKFLGKONDTHSNDKYALAKWGLG 1818
QY 467 -SLAEVKENPPFYGVIRWIGQPPGLNE--VLAGELEDEBACAGCTDGTFRGTRYFTCALX 522
Db 1819 NLLGNAKESERQSITFSV-----ISEGDWQIASAQRKNIAIEKGTSSAHKALAKADR 1873
QY 523 KALFVKLSCRPSRPFASIQPVNSQIERNCSLAFGGYLSV-----VEEN 567
Db 1874 EGLKEVELNRDVAK-----EFINETLIGGTADEAYRSQFIAEHLMTFKMDEN 1922
QY 568 TPKPKEKEGLEIMIGKKGIQGHNSCYLDSTLFCFLPASFSSVLD-----TVLLRPK 618
Db 1923 GBP-IDBKQLEEDINKQ-----FNSVKLKKGFASFKDYWEAYKAIGNIYELR 1970
QY 619 EKNDVEYSETOELLRTIENPLRIYGVYCATIMKLRKILEKVEAASGFTSEK---D 674
Db 1971 EVSD-----QERKMLKTARYTDPET--GKTVKEKIVGVNGLIFNNIQAAKFAAQYVGRFN 2024
QY 675 PBE-----FLNLFPHILRVEPLLKIRSAQKQVQDCYFYQIFMEKNEKVGVPITIOQLE 728
Db 2025 PEKNRVERIYENVYFLH---NPETNGRG-----FSKLPETIAVAFAHKMLE 2066
QY 729 WSFINSNLKFAEAPSLIIO--MPRFGKD 755
Db 2067 GAKIGNKTWIGLSNGLALGNIMEDYCKD 2095

RESULT 8
S38177
SSV7 protein homolog - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein YKR098C
C:Species: Saccharomyces cerevisiae
C>Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 09-Jul-2004
C:Accession: S38177
R:Gaillon, L.; Dujon, B.
A:Submitted to the Protein Sequence Database, March 1994
A:Reference number: S38175
A:Accession: S38177
A:Molecule type: DNA
A:Residues: 1-717 <GAI>
A:Cross-references: UNIPROT:P36026; EMBL:Z28323; NID:G486592; PID:G486593; MIPS:YKR098C
A:Experimental source: strain S288C
C:Genetics:
A:Gene: SGD:UBP11
A:Cross-references: SGD:S0001806; MIPS:YKR098C
A:Map position: 11R

Query Match 2.6%; Score 130; DB 2; Length 717;
Best Local Similarity 23.9%; Pred. No. 0.85; Indels 126; Gaps 18;
Matches 84; Conservative 44; Mismatches 98; Indels 126; Gaps 18;
QY 392 SPPLQPPVNV--SLTTENRF---HSLPFLTKMPNTNGSIGHSPLSLSAQSVMEELNTA 445
Db 138 SLPLRLPLVHMASLSLTHNKFDGSLHETPNELTKPTNDN-----SKEDIVRESNQI 187
QY 446 PVQESPLAMPNGNHSGLVGSIAEVKENPPFYGVIRWIGQPPGLNEVLAGELEDEBAG 505
Db 188 -----ASSNKLEAGS-----EV-----199
QY 506 CTDGTRGTRYFTC--ALKKALFVKLSKCRPSRPFASLQ-----PVSNQIRCN 552
Db 200 -----AYTTSKEALSXPYKL--STGKDALFKTLSSPATAPPVHSLVSSQI--RDS 248
QY 553 SLAFGYLSVVEENTPPKMEK---EGLEIWKKK-----KGIQGHY 591

Db 249 SQDSSSLSKVEK-----PKEEKGKIEAESSAPKAYNLPVIEDSNLLSELITGLQNPC 304
QY 592 NSCYLSTDLFCLEAFASVLDVTLRLPKB--KNDVEYSETQELLRTIENVLEIRY----G 645
Db 305 NTCYINSIIQCLFGTTLFRDLFTLKRYRLFLNTNPKYKEVQ--LSRSIYVLFKKMGLNGR 363
QY 646 YVCATKIMKLRKILEKVEAASGFTSEBKDPDEFRLNLI--PHHILRVEPLK 694
Db 364 AIIPNRLPKCK---KLRPDLNIPDDQODTQEFLLIVLARIHEELSNENVVK 412

RESULT 9
T13827
kinesin-73 - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C:Accession: T13827
R:Li, H.P.; Liu, Z.M.; Nirenberg, M.
Proc. Natl. Acad. Sci. U.S.A. 94, 1086-1091, 1997
A:Title: Kinesin-73 in the nervous system of Drosophila embryos.
A:Reference number: Z17784; MUID:97188425; PMID:9037010
A:Accession: T13827
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1921 <LIH>
A:Cross-references: UNIPROT:O01349; EMBL:U81788; NID:G1906595; PID:G1906596; PIDN:AAB504
C:Genetics:
A:Cross-references: FlyBase:FBgn0019968

Query Match 2.5%; Score 125.5; DB 2; Length 1921;
Best Local Similarity 18.1%; Pred. No. 8.3;
Matches 120; Conservative 94; Mismatches 227; Indels 221; Gaps 30;
QY 27 ECSTVDKQTKLLKVPKSGISQYIODRSVGHRSIPS---AKGKNQIGLKILEQPHAVLF 83
Db 1321 DCSASDGET-----YIEKYTRGVSAVESILTDLRLQNVAVKELETAHQPL 1367
QY 84 VDEDVVEINEKFTTELLAITNCEERFSLFKNRNRLSKGLQIDVGCVPKVLRSGE----138
Db 1368 SMKRTVSV--PNFSQOLI-----NKLQIMRFDASWESLNVGRSESFADL 1411
QY 139 -----EKFP-----GVVFRGCPPLAERTVSGIFGVELLEEGRGQGTGVYQ--181
Db 1412 NNSALGNKFTPAGHSPAGAGVIRSRHS-----FGKGSSDDSPGKAF--GIASP 1459
QY 182 -----GKQLFQDCE--CGFVALDKLELIEDDATALESDYAGPGDTMQLVPLEINSR 233
Db 1460 ATSKLGMRTTLHEEPLGGHRSLE---EPEDSYSDSEYAAEYEQRQNKSWATSR 1515
QY 234 VSLKGGTIBSGTVIFCDVLPKGESLGYFVGVDMDNPIGNWDGRFDGVLCSFACVESTIL 293
Db 1516 --LTAKTMDS---FMDV--SSHNSQSYL-----SVTSSANAM 1547
QY 294 LHINDIPESVQERRPPKLAFMRSRGVGDGKSSSHN-----KPKATGSTD-----339
Db 1548 KHLTGLATLSMSSS-----TSSGYGSAVSCNNLSNEDIASMRSMSIDETPDRVN 1599
QY 340 ---PGNRSEL---FYTLNGSSVDSQPSKSKNTWYIDEVAEDPAKSL----TELSTDF 389
Db 1600 SNSPPNQRARVNPFLKMPKAKIOEQPEPAKK--LQEAFTPLEQLERENASDDDE 1656
QY 390 RSSPPL-----OPPPVNSLTENRHFSLPFLTKM--PNTNGSIGHSPLSLSAQ 437
Db 1657 CAQLPKNNNNVDLVNPKPLSGQMELEEPSQESQTEFATDNQNGNRSDLSHSSD 1716
QY 438 VMEELNAPVQESPL-----AMPNCSHGLE-----464
Db 1717 LLE--GGGIVREELPAGKVVVRKKSNTPQPPSNGNSINNNNGTTQVPRINHRASVAKMEG 1774
QY 465 -----VGSIAEVKENP-----PF--YGVIRWIGQPPGLNEV 493
Db 1775 LAAYLDSSIMTSSTEVDEESKDELVLPELWVGESVLIRPVNTSGVIRFVGTTFQPCA 1834

Qy 494 LAGLELEDECACTDGTFRGTGTYFTCALFKALFVKLKSRRDPSPASLQPVSNQIERCNS 553
Db 1835 WIGVEL-DTPTGKNDGSSVKGQVQYFOCKPKHGMFVRSDKMLMDLRGKAMR-AYKAAEKSNS 1892
Qy 554 LA 555
Db 1893 IS 1894
RESULT 10
F83853
5-methyltetrahydrofolate S-homocysteine methyltransferase meth [imported] - Bacillus hal
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C;Accession: F83853
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000.
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: F83853
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1146 <STO>
A;Cross-references: UNIPROT:Q9KCE1; GB:AP001512; GB:BA000004; NID:g10174030; PIDN:BA8053
A;Experimental source: strain C-125
C;Genetics:
A;Gene: meth
C;Superfamily: cobalamin-dependent methionine synthase; cobalamin-binding homology

Query Match 2.5%; Score 124.5; DB 2; Length 1146;
Best Local Similarity 17.8%; Pred. No. 4.3;
Matches 153; Conservative 115; Mismatches 303; Indels 287; Gaps 40;

Qy 52 DRSVGHSRIPSAKGGK--NOIGL-----KILEQPH-----AVLFVDDVWEINEKF 95
Db 413 DEKVIEQALTYSGKAIINSINLEDGEERFEKVPLVHKYGAADVGTIDEEGMAIT--- 469
Qy 96 TELLALATNCEERFSLFKNNRLSKG-----LOIDVGCVPVKQVLRSEEEKPPGVVRRPG 149
Db 470 AEKLAIV--AKRSYDLVNNKYNIRPSPDIIFDPLFPVGTGDEQYIGSANETVEGIRRIKE 527
Qy 150 PL--LAERTVSGIFPGVELLEBGRGGFTDGVVQKGQLFOCDE---DCGFVALDKLE-- 201
Db 528 ELPECUILLGVSNVSGPLPPV---GREVLNAY-----LYHCTQAGLDYAIWNTBKLERY 579
Qy 202 -----LIEDDDTAL---ESDYAGPGDTMQVELPPLPLEINSRVLKGGTETES 244
Db 580 ASISDEEKELSRKLLFETTDTLAEFTAFYRGKKAKKVETSNLTLEERLA---NYIVE 635
Qy 245 GTVIFCDVLPGKESLGYFVGVDMDNPIGNDGRFD---GVLCSFACVESTILLHNDIIP 301
Db 636 GS-----KDGL-----TEDLDKALAKYDDPLDIINGPLMN-GMDEVGRLFNNNELIV 681
Qy 302 ESVTQERRPPKLAFMGRGVGDKSSSHNKKPKATGSTDPGNRRSELFTYTLNGSSVDSQPQ 361
Db 682 AEVLQSAEVM-----KASVAHLEHMEKKADHDHGKKI-ILATVKGD----- 722
Qy 362 SKSKNTWYIDEVAEDPAKSLTEISTDPRSPPLQPPVPVNSLTTFENRHSFLPSTKMPN 421
Db 723 -----VHDIGKNLVEI-----ILSNNGFRIVNLGIKVTSN 752
Qy 422 -----TNGSIGHSPLSQAQSVMBELNTAPVQESPPLAMPNGSHGLEVGSLA--- 469
Db 753 ELIEAVARENPAIGLS--GLLVKSAQMWLTADDLKQQQISIP-----ILVGGAAALTR 804
Qy 470 ---EVKENPPFYGVIRWIGQPPGLNEVLAGLELEDECACTDGTFRGTGTYFTCALKKALF 526
Db 805 KFTNTKIAEYDGLVY-----AKDAMNGLELAN----- 833
Qy 527 VKLKSRRPDSR---FASLQPVSNQIERCNSIAFGG-----YLSEWVEENTPPKME 573
Db 834 ---KLMKPDREKLAVLSLHKAKEQANRSTONGGGGTAVAKPFRSHVSTTVVAVPPDVK 890

Qy	574	KE-----GLIMTKKKKGIQGHVNSCYLDTSLFCLPAPSSVLDTVLLRPKE	619
Dy	891	PHILRHSHIAHLEPYINMQLRHLGLOKVSRLLAEKDHAL-ELKEKVDALLTRVKE	949
Qy	620	K-----NDVEYYSST-----QELLRTTEIVNP	640
Dy	950	EQLMEAHGMVQFFPAOSGDGDIVDQTGTNEIERPFPNPKPEYLCLADFURPVSSGE	1009
Qy	641	LRIYGVCATKIMKLKILEKVEAASGFTSEEKDPBEFLNILPHILRUEPLLKIRSAQG	700
Dy	1010	MDVVGFLAVTAGKIGBELGEQAEGDY-----LFSLHIQATALEWAEGFAE	1056
Qy	701	KVDCVFYQIFMEKNBKVGPV-----TIQOLLBWSFINSLKPAEPAPSCLIIQMPPRGKD	755
Dy	1057	RV-----HQLM--RDKWGFPSADFTMBERFAAKYGRVRSFG-YPAAC-----POLDDQQ	1102
Qy	756	FKLPKKIFF-SLELNITD	772
Dy	1103	AKLFLKLLKFKGIETLE	1120
RESULT 11			
AD0835			
large repetitive protein [imported] - Salmonella enterica subsp. enterica serovar C;Species: Salmonella enterica subsp. enterica serovar Typhi			
C;Note: This species has also been called Salmonella typhi			
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002			
C;Accession: AD0835			
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.;			
Th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.			
S.; Moule, S.; O'Gaora, P.			
Nature 413, 848-852, 2001			
A;Authors: Parry, C.; Ouail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Ste			
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enter			
A;Reference number: AB0502; UID:21534947; PMID:11677608			
A;Accession: AD0835			
A;Status: preliminary			
A;Molecule type: DNA			
A;Residues: 1-3624 <PAR>			
A;Cross-references: GB:AL513382; PIDN:CAD05867.1; PID:g16503842; GSPDB:GN00176			
C;Genetics:			
A;Gene: STY2875			
Query Match 2.5%; Score 124; DB 2; Length 3624;			
Best Local Similarity 21.9%; Pred. No. 29;			
Matches 103; Conservative 61; Mismatches 167; Indels 140; Gaps 28;			
Qy	191	DCGFVALDKLEIEDDDTALESYAGPG---DTMQVELPPLEINSRVSL-----KGGET	241
Dy	1532	DSGLPAPVAITAEDDVGSIQGINIAGGATDDTMPTLRGTIDIGSTEVEFDIGDSAGFAT	159191
Qy	242	IE-SGVTFICDVLPGKESLYFVGVDMDNFIGNWDGRPD---GVLCFSACVESTILLIH	297
Dy	1592	VDASGNWIFPIATPLSESTHYFT-VQATN--ANGPGGLSAPVGITVDLSAPAQPVSATS	1648
Qy	298	DIIP-----SVTQEREPKLAFMISRGVDKGSSSH---NKPATGSTSD-PGNRRS	345
Dy	1649	DDVPGMTLTLDNGALTINDSRP-----TLNGTGAGATIRILDNGVEIGSATVDOSGNWR-	1702
Qy	346	ELFYTLNGSSVDSPQSCKNTWIIDEVAEDPAKSLTEISTDF----DRSPPLQPPPVN	401
Dy	1703	--FTP-N-TPLES-----NAHIFTAVATDPAGNSQSLSDGGFTLNIDAQAAPDV--PVIT	1749
Qy	402	SLTTENRFHSLPF-----SLTKMPNTNGSIGHSPLSLASQSVMEELMTAPVQESPLAMP	456
Dy	1750	SVIDDNNQTPVPLPGOSTDDROPILNGT-GEPCATTIFDNGTPLGTAAQGENGSWTFF	1808
Qy	457	-PGN-----SHGLEV-----GSLAEVKE-----NPPFYGYVI-RWIGOPPGLNEVL	496
Dy	1809	VPRMLSEGSHNLTVSATDDPAGNTSAVSAPWTIVDITPPALPVLTSVVDDQPGITCNLV	1868
Qy	497	LELBDEC-----AGCT-----DGTF-----RCTRYPTC	519

Db 1869 GOLTNDATPTLNGRGEAGATINVLNPNASIGTTTNSDGTWSFTPTQPLANGSHFTT- 1927

QY 520 ALKKALFVKLSCRPDRSFASIQPVSNQIQRNCNSLAFGGVLSVVBENTPP 570

Db 1928 -----LSATDPAGN-----SSAVSSGFVLTII--DTTPP 1953

RESULT 12

S63244

BN11 protein - yeast (Saccharomyces cerevisiae)

N:Alternate names: protein N0646; protein YNL271C

C:Species: Saccharomyces cerevisiae

C>Date: 27-Apr-1996 #sequence revision 03-May-1996 #text_change 09-Jul-2004

C:Accession: S63244; S63245; S48523; S60909; S65111

R:Sen-Gupta, M.; Lyck, R.; Niedenthal, R.K.; Fleig, U.N.; Hegemann, J.H.

A:Submitted to the Protein Sequence Database, April 1996

A:Reference number: S63235

A:Accession: S63244

A:Molecule type: DNA

A:Residues: 1-1553 <SEN>

A:Cross-references: UNIPROT:P41832; EMBL:Z71547; MIPS:YNL271C

A:Experimental source: strain S288C

R:Messeguy, F.; Dubois, E.; Vierendeels, F.; Scherens, B.; Pierard, A.; Glansdorff, N.

A:Submitted to the Protein Sequence Database, April 1996

A:Reference number: S63245

A:Accession: S63245

A:Molecule type: DNA

A:Residues: 987-1953 <MES>

A:Cross-references: EMBL:Z71547; MIPS:YNL271C

A:Experimental source: strain S288C

R:Fares, H.F.; Pringle, J.R.

A:Submitted to the EMBL Data Library, April 1994

A:Description: Synthetic Lethals of CDC12.

A:Reference number: S48524

A:Accession: S48523

A:Molecule type: DNA

A:Residues: 1-937, 'A', 939-1429, 'C', 1431-1953 <FAR>

A:Cross-references: EMBL:L31766; NID:G472524; PIDN:AAA34455.1; PID:G472525

R:Sen-Gupta, M.; Lyck, R.; Fleig, U.; Niedenthal, R.N.; Hegemann, J.M.

A:Submitted to the EMBL Data Library, October 1995

A:Description: The sequence of a 24152 bp segment from the left arm of chromosome XIV from Sa

A:Reference number: S60909

A:Accession: S60909

A:Molecule type: DNA

A:Residues: 1-1553 <SE2>

A:Cross-references: EMBL:X92494; NID:G1045236; PIDN:CAA63225.1; PID:G1045237

R:Sen-Gupta, M.; Lyck, R.; Fleig, U.; Niedenthal, R.K.; Hegemann, J.H.

Yeast 12, 505-514, 1996

A>Title: The sequence of a 24 152 bp segment from the left arm of chromosome XIV from Sa

A:Reference number: S65111; MUID:96310631; PMID:8740425

A:Accession: S65111

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-1553 <SEN>

A:Cross-references: EMBL:X92494; NID:G1045236; PIDN:CAA63225.1; PID:G1045237

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1995

C:Genetics:

A:Gene: SGD:BN11; SHE5; SYL39

A:Cross-references: SGD:S0005215; MIPS:YNL271C

A:Map position: 14L

Query Match

Best Local Similarity 19.3%; Pred. No. 14;

Matches 181; Conservative 100; Mismatches 300; Indels 357; Gaps 46;

QY 57 HSRIPSAKGNQIGLKLEQHAVLFVDEDDVVEINE---KFTLELLAIT-----NCEER 108

Db 902 HIKKPKVKQKMKR-DRKPLVKPQ-----EADVKNLDRALTEIQMESNDISKFNVEER 955

QY 109 FS-LFKNRNLRSKGLQIDVQCPVKVQSRGEEKFPVGRFLLAERTVSGIFFGVELL 167

Db 956 VNELFNEKSLA-----LKRLELETKYKGF-----GIDFNVDIEI 990

QY 168 EEGRGQGTDDVYQKQLFOCEDDCGFVALD-----KL 200

Db 991 MDSFKNTGD-----VETEDANTASLDPTYQKKLDEINRITDOLLDIQTQTEHEI 1042

QY 201 ELTEDDDTALESYAG--PGDTMQVELPPELEINRSVS-LKGGETIESGTVIFCDVLPQKE 257

Db 1043 QVEDGESDLSSSSDDESEIYQDASPTQLRSEHSELSSG-----SGPGSFLDALSQKY 1098

QY 258 SLGYFVGVDV---DNPIGNWDRFPGVLCFACVESTILHINDIIPESVTOERRPPKLA 314

Db 1099 GTGQNVTAASAAFGNNNGS-----GIGPLHKSVEKTFM---NRLRSTVSS--APYLE 1146

QY 315 FMERGIV-----DKGSSSHKPKATGSTSDPGNRRSELFTYLANG----- 353

Db 1147 ELTKQKNVPEYQNEDEGLDKSLPENSTASAAAFDKAKDMRQ--HVENKQGRVVNH 1205

QY 354 ---SSVDSQPQSKNTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPV-----NS 402

Db 1206 EEDKTADFSAVSKLNT-----DGAEDLSTQSSVLSQPPPPPPPPVPAKLFGES 1257

QY 403 LTTENRPHSLPFLTKMPNTNGSIGHSPLSLSAQSVMEELNTPVQESPPPLAMPPGNSHG 462

Db 1258 LEKEKKSSED--DTVKQETT-----GDSP-----APPPPPPPPPPPPPMALFG 1296

QY 463 LEVGSALAEVKENPFYGVIRWIGOPPCGLNEVLAG----- 496

Db 1297 KPKG-----ETPP-----PPPLPSVSSSTDGVIAPPMPMPASQIKSAVTSP 1340

QY 497 -----LELEDECAGCTDGTFRGT-----RVFT 518

Db 1341 LPQSPSLFEKYPRPHKKLKQLHWEKLDCTNSIWTGKAEFADDLYEKGVLADLEKAPA 1400

QY 519 CALKXALFVKLSCRPDRSFASIQPVSNQ-----IERCNSLAFGGVLSVWV---ENTP 569

Db 1401 AREIKSLASKRKEDLQKITFLS-RDISQQGGINLHMYSSLSVADIVKILNCNDRDFLTGP 1459

QY 570 PKME-----KEGLEIMIGKKGIQGHYNSCYLDSTLFCLEAFSSVLDVTLVLRPKEK--NDV 623

Db 1460 SVVEFLSKSIIIEVSV---NLARNVAPYSTD-----WEGVRNLEDAKPEKDPNDL 1507

QY 624 E-----YYSETQELLRTIEIYNPLRIYGVCAATKIMKLKILEKVEA----- 664

Db 1508 QRADQIYQLMVNLESYWGSRMAL-TVVTYSEREYNELLA---KLRKVDKAVSALOES 1562

QY 665 -----ASG-----FTSEKDPPEEFLNII----- 682

Db 1563 DNLNVNPNVILAVGNFNWDTSKQAGFKLSTQLRLTFIKDTTNSMTFLNYVEKIVRLNYP 1622

QY 683 -FHHL-RVEPLLKI--RSAGQKQDCYFYQIFMEKNEKVGVTPTIQQLLESFINSNLKF 738

Db 1623 SFNDFLSELEPVLVDVVKVSIQVLVNDG-----KQFSQSIIVNVERVEIGNLSDSKF 1674

QY 739 AEAPSLIIQMPRFGDKFLKFIKFPSELEINITDLED 776

Db 1675 HPLDKVLITLTP-----VLPEARKK-GDLED 1700

RESULT 13

T02345

hypothetical protein KIAA0324 - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 09-Jul-2004

C:Accession: T02345

R:Ricke, D.O.; Bruce, D.; Mundt, M.; Doggett, N.; Munk, C.; Saunders, E.; Robinson, D.; re, J.; White, S.; Ueng, S.; Tatum, O.; Campbell, C.; Fawcett, J.; Deaven, L.

submitted to the EMBL Data Library, March 1998

A:Description: Sequencing of human chromosome 16p13.3.

A:Reference number: Z14664

A:Accession: T02345

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1791 <RIC>

A:Cross-references: UNIPROT:O60382; EMBL:AC004493; NID:G2996648; PIDN:AA08453.1; PID:G2

308 SDICTSLPQNSQIQQMHPQNI GS--DSNSFSNLA VGVKSESSPQGVPSKQENTIMS 366
 421 NTNGSTIGH-----SPLSLSAQSYMEELNAPVQESPLMAPPPGNHGLE- 464
 367 NAISSGKHIOEDPRRIITGWDEAQPNNLTGSGVIGQHTTISE-----SHNLQN 416
 465 -VGS-----LAEVKENPPFYGVIRWI-----GQPLGLNEVLAGLEHDECA- 504
 417 SIGTTCRYGNVSHDPKFKQKQWLLFLRHARSKCPFG-----GRCDQNCVTVQKLWSH 470
 505 --GCTDGTGRGTRYFTCA LKKALFVKLKSCRDPDSRFASLPVSNQTERCN----- 552
 471 MDMCAPOQ---CLYPRCRHTKALIGHYKNCK-DPRCPVCVPVTVYQQQANRALARKNE 526
 553 SLAFGGYLSEWVE-----ENTPPKMEKEGLEIMTKKKGIQGHY 591
 527 SSAVGSVNRSVVNSDLSANAGAVSCTPRCADTLDNLP SLKRLKVE-----QSFQP-- 578
 592 NSCYLSTUFLCLPAFSSVLDTVLLRPEKNDVB-YYSETQELLRLTELVNPLRIYGVYCAT 650
 579 ---VWPKTESCKSSIVSTTEADLSQDAERKDRHPLKSETME-VKVEIPLD----- 623
 651 KIMKLARKILEKVEAASGFTSEBKDPBEFLNLFPHHLIRVEPLKIRKSAGKVQDCYFQI 710
 624 -----NSVQAGFIKETKEEP-----FENVPKPKVPSEPGKHGUSGDSPKQENI 667
 711 FM EK-----NEKVGVPVTIOQLLEWSFINSNLKFAEAPSLIIQMPRFGKDFLKFKI FPS 765
 668 KMKKEPGWPKKEGCPCKBELVSEPLTS-----KSRKPK 702
 766 LB-LNITD LLEDTPQCR-----ICGLLWYEGREC-----Y 796
 703 IKGVS LT ELP--TPEQVREHIRGLRWVGSGKAKAEKNQAMENSMSENSCQLCAVEKLTF 760
 797 DDPDISAGKIKQCKYKTCNTQVHLLHPKLNHKNYPSLPKDLPDWDWRH-CCIPCONMELF 855
 761 EPPPI-----YCTPGARI-----KRNAMYTVTG-----GGETHYFCIPICYNESRG 802
 856 AVL CIE-TSHYAVFYKVYKDD-----SAWLFFDSM-----ADRDGGQGNFIP 897
 803 DTILAEGTSPKAKLEKKNDDEIESWVQCDKQAWHQICALFNGERNDDGGAEYTCP 862
 RESULT 15
 T10955
 early nodulin binding protein 1 - spring vetch
 C;Species: *Vicia sativa* (spring vetch, fava)
 C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
 C;Accession: T10955
 R;Christiansen, A.; Hansen, A.C.; Vijn, I.; Pallisgaard, N.; Larsen, K.; Yang, W.
 submitted to the EMBL Data Library, December 1995
 A;Description: A novel type of DNA binding protein interacts with a conserved s
 A;Reference number: 217228
 A;Accession: T10955
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-1641 <CHR>
 A;Cross-references: UNIPROT:Q41700; EMBL:X95995; NID:gl3360633; PID:e225826
 C;Genetics:
 A;Note: ENBP1
 C;Keywords: DNA binding
 Query Match 2.4%; Score 121; DB 2; Length 1641;
 Best Local Similarity 19.4%; Pred. No. 13;
 Matches 210; Conservative 109; Mismatches 311; Indels 452; Gaps 57
 107 ERFSLFKRNRLSKGLQIDVGC PKVKQLRSGE-----EKPGP-VYFRFGPPLAERTVSGI 160
 100 ESFNLWQE---GEGQQVEV-----FEGSGNLKGLFDGVEFLGGFVEDNRNVGL 146
 161 -----FRGVELLEBGRGGGFTDGYVQKQLFCQDEDCGFVALDKLELIEDDDTALES 212
 147 GQPMSSVGVFG-----NAGGVAGYKGDHGRVDGVCG-----NDLSAPHS 187

Search completed: April 18, 2005, 15:25:28
Job time : 28 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 15:22:19 ; Search time 355 Seconds
(without alignments)
3122.360 Million cell updates/sec

Title: US-09-671-687A-3

Perfect score: 5034

Sequence: 1 MSSGLWSQEKVTSFYWEERI.....RLLCDAYMCVQSPMTSLYK 949

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 6959266 seqs, 1168006243 residues

Total number of hits satisfying chosen parameters: 6959266

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents AA Main:*

- 1: /cgn2_6/ptodata/1/paa/PCTUS COMB.pcp.*
- 2: /cgn2_6/ptodata/1/paa/US06 COMB.pcp.*
- 3: /cgn2_6/ptodata/1/paa/US07 COMB.pcp.*
- 4: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 5: /cgn2_6/ptodata/1/paa/US081 COMB.pcp.*
- 6: /cgn2_6/ptodata/1/paa/US082 COMB.pcp.*
- 7: /cgn2_6/ptodata/1/paa/US083 COMB.pcp.*
- 8: /cgn2_6/ptodata/1/paa/US084 COMB.pcp.*
- 9: /cgn2_6/ptodata/1/paa/US085 COMB.pcp.*
- 10: /cgn2_6/ptodata/1/paa/US086 COMB.pcp.*
- 11: /cgn2_6/ptodata/1/paa/US087 COMB.pcp.*
- 12: /cgn2_6/ptodata/1/paa/US088 COMB.pcp.*
- 13: /cgn2_6/ptodata/1/paa/US089 COMB.pcp.*
- 14: /cgn2_6/ptodata/1/paa/US090 COMB.pcp.*
- 15: /cgn2_6/ptodata/1/paa/US091 COMB.pcp.*
- 16: /cgn2_6/ptodata/1/paa/US092 COMB.pcp.*
- 17: /cgn2_6/ptodata/1/paa/US093 COMB.pcp.*
- 18: /cgn2_6/ptodata/1/paa/US094 COMB.pcp.*
- 19: /cgn2_6/ptodata/1/paa/US095 COMB.pcp.*
- 20: /cgn2_6/ptodata/1/paa/US096 COMB.pcp.*
- 21: /cgn2_6/ptodata/1/paa/US097A COMB.pcp.*
- 22: /cgn2_6/ptodata/1/paa/US097B COMB.pcp.*
- 23: /cgn2_6/ptodata/1/paa/US098 COMB.pcp.*
- 24: /cgn2_6/ptodata/1/paa/US099A COMB.pcp.*
- 25: /cgn2_6/ptodata/1/paa/US099B COMB.pcp.*
- 26: /cgn2_6/ptodata/1/paa/US100 COMB.pcp.*
- 27: /cgn2_6/ptodata/1/paa/US101 COMB.pcp.*
- 28: /cgn2_6/ptodata/1/paa/US102 COMB.pcp.*
- 29: /cgn2_6/ptodata/1/paa/US103 COMB.pcp.*
- 30: /cgn2_6/ptodata/1/paa/US104 COMB.pcp.*
- 31: /cgn2_6/ptodata/1/paa/US105 COMB.pcp.*
- 32: /cgn2_6/ptodata/1/paa/US106 COMB.pcp.*
- 33: /cgn2_6/ptodata/1/paa/US107 COMB.pcp.*
- 34: /cgn2_6/ptodata/1/paa/US108 COMB.pcp.*
- 35: /cgn2_6/ptodata/1/paa/US109 COMB.pcp.*
- 36: /cgn2_6/ptodata/1/paa/US110 COMB.pcp.*
- 37: /cgn2_6/ptodata/1/paa/US60 COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

RESULT 1
US-09-671-687A-3
; Sequence 3, Appli
; GENERAL INFORMATION:
; APPLICANT: WALLACH, David
; APPLICANT: KOVALENKO, Andrei
; APPLICANT: CANTARELLA, Giuseppe
; TITLE OF INVENTION: INHIBITOR OF NF-KB ACTIVATOR
; FILE REFERENCE: WALLACH=25
; CURRENT APPLICATION NUMBER: US/09/671,687A
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 09/646,403
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: IL 126024
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: IL 134604
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1

ALIGNMENTS

Result No.	Score	Query Match	Length	DB	ID	Description
1	5034	100.0	949	20	US-09-671-687A-3	Sequence 3, Appli
2	4983	99.0	953	1	PCT-US02-27777-137	Sequence 137, App
3	4983	99.0	953	1	PCT-US02-27777A-137	Sequence 137, App
4	4983	99.0	953	27	US-10-170-205B-37570	Sequence 37570, A
5	4983	99.0	960	1	PCT-US04-07268-250	Sequence 250, App
6	4983	99.0	960	33	US-10-788-792-250	Sequence 250, App
7	4971.5	98.8	956	1	PCT-US02-14570-4	Sequence 4, Appli
8	4971.5	98.8	956	23	US-09-851-673-4	Sequence 4, Appli
9	4971.5	98.8	956	33	US-10-755-889-490	Sequence 490, App
10	4971.5	98.8	956	37	US-60-440-068-490	Sequence 490, App
11	4971.5	98.8	956	37	US-60-469-757-490	Sequence 490, App
12	3876.5	77.0	739	1	PCT-US01-01239-1743	Sequence 1743, Ap
13	3876.5	77.0	739	22	US-09-764-902-1743	Sequence 1743, Ap
14	3862	76.7	731	22	US-09-786-797B-9	Sequence 9, Appli
15	3862	76.7	731	35	US-10-921-707-9	Sequence 9, Appli
16	3862	76.7	731	37	US-60-131-321-7	Sequence 7, Appli
17	3855	76.6	731	20	US-09-629-469A-18843	Sequence 18843, A
18	3855	76.6	731	35	US-10-917-503-18843	Sequence 18843, A
19	3636	72.2	685	18	US-09-488-725A-2399	Sequence 2399, Ap
20	3636	72.2	685	28	US-10-258-898A-2399	Sequence 2399, Ap
21	3636	72.2	685	28	US-10-286-897-2399	Sequence 2399, Ap
22	3563	70.8	698	18	US-09-488-725A-5971	Sequence 5971, Ap
23	3563	70.8	698	28	US-10-258-898A-5971	Sequence 5971, Ap
24	3563	70.8	698	28	US-10-286-897-5971	Sequence 5971, Ap
25	2755	54.7	558	1	PCT-US01-01239-1207	Sequence 1207, Ap
26	2755	54.7	558	22	US-09-764-902-1207	Sequence 1207, Ap
27	2455	48.8	512	1	PCT-US01-14827-12400	Sequence 12400, A
28	2343	46.5	476	20	US-09-629-469A-18587	Sequence 18587, A
29	2343	46.5	476	35	US-10-917-503-18587	Sequence 18587, A
30	1751	34.8	394	1	PCT-US01-14827-12399	Sequence 12399, A
31	1275	25.3	261	1	PCT-US01-16450-1609	Sequence 1609, Ap
32	1275	25.3	261	1	PCT-US01-16450A-1609	Sequence 1609, Ap
33	1275	25.3	261	28	US-10-264-237-1609	Sequence 1609, Ap
34	1173	23.3	238	1	PCT-US01-16450A-1610	Sequence 1610, Ap
35	1173	23.3	238	1	PCT-US01-16450A-1610	Sequence 1610, Ap
36	1173	23.3	238	28	US-10-264-237-1610	Sequence 1610, Ap
37	856	17.0	517	20	US-09-614-150-11799	Sequence 11799, A
38	856	17.0	517	20	US-09-614-150A-11799	Sequence 11799, A
39	856	17.0	517	37	US-60-191-637-11831	Sequence 11831, A
40	856	17.0	517	37	US-60-191-681-9275	Sequence 9275, Ap
41	786.5	15.6	167	1	PCT-US01-14827-15309	Sequence 15309, A
42	775	15.4	188	22	US-09-758-472-8763	Sequence 8763, Ap
43	775	15.4	188	28	US-10-235-926-8763	Sequence 8763, Ap
44	742.5	14.7	583	37	US-60-167-217-11813	Sequence 11813, A
45	742.5	14.7	583	37	US-60-173-464-9570	Sequence 9570, Ap

; SEQ ID NO 3
; LENGTH: 949
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-671-687a-3

Query Match 100.0%; Score 5034; DB 20; Length 949;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 949; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSGLMSQEKVTSFYWEERIFVYLLLOECSTVDKOTKLLKVPKSGIGQYIQDRSVGHRI 60
DB 1 MSSGLMSQEKVTSFYWEERIFVYLLLOECSTVDKOTKLLKVPKSGIGQYIQDRSVGHRI 60

QY 61 PSAKGKKNQIGLKILQPHAVLFVDEDDVVEINEKFTTELLAINTCERFSLFKNRRLSK 120
DB 61 PSAKGKKNQIGLKILQPHAVLFVDEDDVVEINEKFTTELLAINTCERFSLFKNRRLSK 120

QY 121 GLQIDVGCPCVKQLRSGBEKFPGVVRFRGPLLAERTVSGIFFGVLELLEGRGGQFTDGVY 180
DB 121 GLQIDVGCPCVKQLRSGBEKFPGVVRFRGPLLAERTVSGIFFGVLELLEGRGGQFTDGVY 180

QY 181 QGKQLFOCEDDCGFVALDKLELIEDDDTALESDYAGPGDTMQVELPPELNSRVSLKGG 240
DB 181 QGKQLFOCEDDCGFVALDKLELIEDDDTALESDYAGPGDTMQVELPPELNSRVSLKGG 240

QY 241 TTIESGTVIFCDVLPKESLGYPVGVDMNDPIGNWDRFGDVLCSFACVESTILLHNDII 300
DB 241 TTIESGTVIFCDVLPKESLGYPVGVDMNDPIGNWDRFGDVLCSFACVESTILLHNDII 300

QY 301 PESVTQERRPPKLAFMRSRGVGDGSSSHNKPATGSTSDPGNRRSELFTYTLNGSSVDSQP 360
DB 301 PESVTQERRPPKLAFMRSRGVGDGSSSHNKPATGSTSDPGNRRSELFTYTLNGSSVDSQP 360

QY 361 QSKSKNTWTIDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSLTNNRSHSLPFSLTMP 420
DB 361 QSKSKNTWTIDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSLTNNRSHSLPFSLTMP 420

QY 421 NTNGSIHSPSLSLSAOSVMEELNTAPVQSSPPLAMPNGSHGLEVSLAEVKNPFFYGV 480
DB 421 NTNGSIHSPSLSLSAOSVMEELNTAPVQSSPPLAMPNGSHGLEVSLAEVKNPFFYGV 480

QY 481 IRWIGOPPGNEVLAGELEDEDECACTDGTFRGTRFTCALKKALFKVLSKCRPDSRFAS 540
DB 481 IRWIGOPPGNEVLAGELEDEDECACTDGTFRGTRFTCALKKALFKVLSKCRPDSRFAS 540

QY 541 LQVSNQIERCNSLAFGGVYLSEVVEENTPPKMEKEGLEIMIGKKGIQGHYNSCYLDSTL 600
DB 541 LQVSNQIERCNSLAFGGVYLSEVVEENTPPKMEKEGLEIMIGKKGIQGHYNSCYLDSTL 600

QY 601 FCLFAFSSVLDVTLRPKEKNDVEYSETQELLARTEIVNPLRIYGVYCATKIMKLRKILE 660
DB 601 FCLFAFSSVLDVTLRPKEKNDVEYSETQELLARTEIVNPLRIYGVYCATKIMKLRKILE 660

QY 661 KVEAASGFTSEEDKDPREFNLPHILRVBPLKIRSAQOKVQDQCFYQIFMEKNEKVG 720
DB 661 KVEAASGFTSEEDKDPREFNLPHILRVBPLKIRSAQOKVQDQCFYQIFMEKNEKVG 720

QY 721 PTIQQLLEWSFINSLKFAEAPASCLIIQMPFRGDKFKKIPPSLELNTDLEDTPRQ 780
DB 721 PTIQQLLEWSFINSLKFAEAPASCLIIQMPFRGDKFKKIPPSLELNTDLEDTPRQ 780

QY 781 CRTCGGLAMECECYDDPDISAGTKQFCKTNTQVHLHPKRLNHKNYVSLPKOLPDW 840
DB 781 CRTCGGLAMECECYDDPDISAGTKQFCKTNTQVHLHPKRLNHKNYVSLPKOLPDW 840

QY 841 DWRHGICPCONMELFVLCITESHYAFVKYGDSDSAWLPFFDSMADRDGGQNGFNIPQVT 900
DB 841 DWRHGICPCONMELFVLCITESHYAFVKYGDSDSAWLPFFDSMADRDGGQNGFNIPQVT 900

QY 901 PCPEVGYLKWLEDLHSLDSRRIQGCARLLCDAYNCMTQSPMTSLYK 949
DB 901 PCPEVGYLKWLEDLHSLDSRRIQGCARLLCDAYNCMTQSPMTSLYK 949

RESULT 2
PCT-US02-27777-137
; Sequence 137, Application PC/TUS0227777
; GENERAL INFORMATION:
; APPLICANT: diaDexus, Inc.
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; APPLICANT: Salceda, Susana
; TITLE OF INVENTION: Compositions and Methods Relating to Breast Specific Genes and P
; FILE REFERENCE: DEX-0346
; CURRENT APPLICATION NUMBER: PCT/US02/27777
; CURRENT FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: US 60/316,306
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 137
; LENGTH: 953
; TYPE: PRT
; ORGANISM: Homo sapien
PCT-US02-27777-137

Query Match 99.0%; Score 4983; DB 1; Length 953;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 948; Conservative 0; Mismatches 1; Indels 4; Gaps 4;

QY 1 MSSGLMSQEKVTSFYWEERIFVYLLLOECSTVDKOTKLLKVPKSGIGQYIQDRSVGHRI 60
DB 1 MSSGLMSQEKVTSFYWEERIFVYLLLOECSTVDKOTKLLKVPKSGIGQYIQDRSVGHRI 60

QY 61 PSAKGKKNQIGLKILQPHAVLFVDEDDVVEINEKFTTELLAINTCERFSLFKNRRLS 119
DB 61 PSAKGKKNQIGLKILQPHAVLFVDEDDVVEINEKFTTELLAINTCERFSLFKNRRLS 120

QY 120 KGLQIDVGCPCVKQLRSGBEKFPGVVRFRGPLLAERTVSGIFFGVLELLEGRGGQFTDGV 179
DB 120 KGLQIDVGCPCVKQLRSGBEKFPGVVRFRGPLLAERTVSGIFFGVLELLEGRGGQFTDGV 180

QY 180 YQKQQLFOCEDDCGFVALDKLELIEDDDTALESDYAGPGDTMQVELPPELNSRVSLK 238
DB 180 YQKQQLFOCEDDCGFVALDKLELIEDDDTALESDYAGPGDTMQVELPPELNSRVSLK 240

QY 239 GETIESGTVIFCDVLPKESLGYPVGVDMNDPIGNWDRFGDVLCSFACVESTILLHIN 297
DB 241 GETIESGTVIFCDVLPKESLGYPVGVDMNDPIGNWDRFGDVLCSFACVESTILLHIN 300

QY 298 DIIIPESVTQERRPPKLAFMRSRGVGDGSSSHNKPATGSTSDPGNRRSELFTYTLNGSSV 356
DB 301 DIIIPESVTQERRPPKLAFMRSRGVGDGSSSHNKPATGSTSDPGNRRSELFTYTLNGSSV 360

QY 357 DSQPSQSKNTWTIDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSLTNNRSHSLPFSL 416
DB 361 DSQPSQSKNTWTIDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSLTNNRSHSLPFSL 420

QY 417 TKMPNTNGSIHSPSLSLSAOSVMEELNTAPVQSSPPLAMPNGSHGLEVSLAEVKNP 476
DB 421 TKMPNTNGSIHSPSLSLSAOSVMEELNTAPVQSSPPLAMPNGSHGLEVSLAEVKNP 480

QY 477 FYGVIRWIGOPPGNEVLAGELEDEDECACTDGTFRGTRFTCALKKALFKVLSKCRPDS 536
DB 481 FYGVIRWIGOPPGNEVLAGELEDEDECACTDGTFRGTRFTCALKKALFKVLSKCRPDS 540

QY 537 RFASLQPVSNQIERCNSLAFGGVYLSEVVEENTPPKMEKEGLEIMIGKKGIQGHYNSCYL 596
DB 541 RFASLQPVSNQIERCNSLAFGGVYLSEVVEENTPPKMEKEGLEIMIGKKGIQGHYNSCYL 600

QY 597 DSTLFCLFAFSSVLDVTLRPKEKNDVEYSETQELLARTEIVNPLRIYGVYCATKIMKLR 656
DB 601 DSTLFCLFAFSSVLDVTLRPKEKNDVEYSETQELLARTEIVNPLRIYGVYCATKIMKLR 660

QY 657 KILEKVEAASGFTSEEDKDPREFNLPHILRVBPLKIRSAQOKVQDQCFYQIFMEKNE 716

Db 661 KILEKVEAASGFTSEBKDPSEFLNLFPHLLRVEPLKIRSAQKQVQDCYFYQIFMEKNE 720
QY 717 KVGVPPIQQLLEWSFINSNLKFAEAPSCIIQMPRFQKDFKPKIFPSLELNITDLED 776
Db 721 KVGVPPIQQLLEWSFINSNLKFAEAPSCIIQMPRFQKDFKPKIFPSLELNITDLED 780
QY 777 TPRQCRICGLAMYECCYDDPDISAGKIQKFCCTCNTQVHLHPKRLNHNKYNPVS LPKD 836
Db 781 TPRQCRICGLAMYECCYDDPDISAGKIQKFCCTCNTQVHLHPKRLNHNKYNPVS LPKD 840
QY 837 LPDMDWRHGCIPCONMELFAVLCTIETSHYVAFVKYKGDSDSAWLFDDSMADRDGQNGFNI 896
Db 841 LPDMDWRHGCIPCONMELFAVLCTIETSHYVAFVKYKGDSDSAWLFDDSMADRDGQNGFNI 900
QY 897 PQVTPCPEVGEYLKMSLEDLHSLDSRRIOGCARRLLCDAYMCYQSPMTSLYK 949
Db 901 PQVTPCPEVGEYLKMSLEDLHSLDSRRIOGCARRLLCDAYMCYQSPMTSLYK 953

RESULT 3

PCT-US02-27777A-137
; Sequence 137, Application PC/TUS0227777A
; GENERAL INFORMATION:
; APPLICANT: diadexus, Inc.
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; APPLICANT: Salceda, Susana
; TITLE OF INVENTION: Compositions and Methods related to Breast Specific Genes and Pct
; FILE REFERENCE: DEX-0346
; CURRENT APPLICATION NUMBER: PCT/US02/27777A
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: US 60/316,307
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 137
; LENGTH: 953
; TYPE: PRT
; ORGANISM: Homo sapien
PCT-US02-27777A-137

Query Match 99.0%; Score 4983; DB 1; Length 953;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 948; Conservative 0; Mismatches 1; Indels 4; Gaps 4;

QY 1 MSSGLWSQEKVTSPLYWEERIFYLLLOECSTVDTKOTKLLKVPKGSIGQYIQDRSVGHSRI 60
Db 1 MSSGLWSQEKVTSPLYWEERIFYLLLOECSTVDTKOTKLLKVPKGSIGQYIQDRSVGHSRI 60
QY 61 PSAKGNKQIGLKILQPHAVLFDVE-DVVEINEKETELLAITNCEERFSLFKNRNL 119
Db 61 PSAKGNKQIGLKILQPHAVLFDVEKDVVEINEKETELLAITNCEERFSLFKNRNL 120
QY 120 KGLQIDVGCPCVKQLRSGBEKEFGVVRFRGPLLAERTVSGIFPGVELLEGRGQGTG 179
Db 121 KGLQIDVGCPCVKQLRSGBEKEFGVVRFRGPLLAERTVSGIFPGVELLEGRGQGTG 180
QY 180 YQKQLFOCEDCG-FVALDKLELIEDDDTALESYAGPDTMQVELPPLPINSRVSLKG 238
Db 181 YQKQLFOCEDCGFVVALDKLELIEDDDTALESYAGPDTMQVELPPLPINSRVSLKV 240
QY 239 GETIESGTIVFCVLPKGSLSGYFVGVDMDNPIGNWDRPDGV-LCSFACVESTILLHN 297
Db 241 GETIESGTIVFCVLPKGSLSGYFVGVDMDNPIGNWDRPDGVQVQLCSFACVESTILLHN 300
QY 298 DIIPSVTOERRPPKLAFWRSVGDKGSSHNKPKATGSTSDPGNR-RSELFTYLANGSSV 356
Db 301 DIIPSVTOERRPPKLAFWRSVGDKGSSHNKPKATGSTSDPGNRNRSSELFYTLNGSSV 360
QY 357 DSQPSKSKNTWYIDVBAEDPAKSLTEISTDFRSSPPLQPPVNSLTITENRPHSLPFS 416
Db 361 DSQPSKSKNTWYIDVBAEDPAKSLTEISTDFRSSPPLQPPVNSLTITENRPHSLPFS 420

RESULT 4

US-10-170-205E-37570
; Sequence 37570, Application US/10170205E
; GENERAL INFORMATION:
; APPLICANT: ADAMS, Mark
; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN
; TITLE OF INVENTION: CAPTURE AGENTS, AND USES THEREOF
; FILE REFERENCE: CL001381
; CURRENT APPLICATION NUMBER: US/10/170, 205E
; CURRENT FILING DATE: 2002-06-13
; NUMBER OF SEQ ID NOS: 40312
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 37570
; LENGTH: 953
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-170-205E-37570

Query Match 99.0%; Score 4983; DB 27; Length 953;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 948; Conservative 0; Mismatches 1; Indels 4; Gaps 4;

QY 1 MSSGLWSQEKVTSPLYWEERIFYLLLOECSTVDTKOTKLLKVPKGSIGQYIQDRSVGHSRI 60
Db 1 MSSGLWSQEKVTSPLYWEERIFYLLLOECSTVDTKOTKLLKVPKGSIGQYIQDRSVGHSRI 60
QY 61 PSAKGNKQIGLKILQPHAVLFDVE-DVVEINEKETELLAITNCEERFSLFKNRNL 119
Db 61 PSAKGNKQIGLKILQPHAVLFDVEKDVVEINEKETELLAITNCEERFSLFKNRNL 120
QY 120 KGLQIDVGCPCVKQLRSGBEKEFGVVRFRGPLLAERTVSGIFPGVELLEGRGQGTG 179
Db 121 KGLQIDVGCPCVKQLRSGBEKEFGVVRFRGPLLAERTVSGIFPGVELLEGRGQGTG 180
QY 180 YQKQLFOCEDCG-FVALDKLELIEDDDTALESYAGPDTMQVELPPLPINSRVSLKG 238

```

Db 181 YGKQLFQDCEDCGVFVLDKLEIEDDDTALSDYAGPDTMQVELPLEINSRVSLKV 240
Qy 239 GETIESGTIVFCVLPKESLGYFVGVDMNDPIGNWDRFDGV-LCSFACVESTILLHIN 297
Db 241 GETIESGTIVFCVLPKESLGYFVGVDMNDPIGNWDRFDGV-LCSFACVESTILLHIN 300
Qy 298 DIIPESVTOERRPPKLAFAVRGVDGKSSSHNPKATGSTDGPNR-RSELYFTLNGSSV 356
Db 301 DIIPESVTOERRPPKLAFAVRGVDGKSSSHNPKATGSTDGPNR-RSELYFTLNGSSV 360
Qy 357 DSQPSKSNWTYIDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSLTITENRPHSLPFL 416
Db 361 DSQPSKSNWTYIDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSLTITENRPHSLPFL 420
Qy 417 TKMPNTNGSIGHSPISLSAQSVMEELNAPVQESPPPLAMPNGSHGLEVGSIAEVKENPP 476
Db 421 TKMPNTNGSIGHSPISLSAQSVMEELNAPVQESPPPLAMPNGSHGLEVGSIAEVKENPP 480
Qy 477 FYGVIRWIGQPPGLNEVLAGLEDEBACAGTGTGTRGTRYFTCALKKALFVKLSCRPS 536
Db 481 FYGVIRWIGQPPGLNEVLAGLEDEBACAGTGTGTRGTRYFTCALKKALFVKLSCRPS 540
Qy 537 RFASLOPVSNQIERCNSLAFGGYLSVVEENTPPKMEKEGLEIMIGKKGIQGHYNSCYL 596
Db 541 RFASLOPVSNQIERCNSLAFGGYLSVVEENTPPKMEKEGLEIMIGKKGIQGHYNSCYL 600
Qy 597 DSTLFCFLAFSSVLDTVLLRPKEKNDEYVYSETQELLRTIENVPLRIYGVVCATKIMKLR 656
Db 601 DSTLFCFLAFSSVLDTVLLRPKEKNDEYVYSETQELLRTIENVPLRIYGVVCATKIMKLR 660
Qy 657 KILEKVEAASGFTSEKDPPEEFNLIFHILRVEPLLKIRSAQKQVQDCYFYQIFMEKNE 716
Db 661 KILEKVEAASGFTSEKDPPEEFNLIFHILRVEPLLKIRSAQKQVQDCYFYQIFMEKNE 720
Qy 717 KVGVPITQQLLEWSFINSNLKFAEAPSLIIQMPREGKDFKPKIFPSLELNTDILLED 776
Db 721 KVGVPITQQLLEWSFINSNLKFAEAPSLIIQMPREGKDFKPKIFPSLELNTDILLED 780
Qy 777 TPRQCRICGLAMVYECRECYDDPDISAGKIKQKCTCNTQVHLHPKRLNHNKYNPVSIPKD 836
Db 781 TPRQCRICGLAMVYECRECYDDPDISAGKIKQKCTCNTQVHLHPKRLNHNKYNPVSIPKD 840
Qy 837 LPDMDWRHGCIPCONNELFAVLCTIETSHYVAFVKYKDDSAWLPFDSMADRDGGQGNFI 896
Db 841 LPDMDWRHGCIPCONNELFAVLCTIETSHYVAFVKYKDDSAWLPFDSMADRDGGQGNFI 900
Qy 897 PQVTPCPEVGEYLKMSLEDLHSLDSRRIOGCARRLLCDAYMCMYQSPMTSLYK 949
Db 901 PQVTPCPEVGEYLKMSLEDLHSLDSRRIOGCARRLLCDAYMCMYQSPMTSLYK 953

```

RESULT 5

```

PCT-US04-07268-250
; Sequence 250, Application PC/TUS0407268
; GENERAL INFORMATION:
; APPLICANT: Bayer Pharmaceuticals Corporation
; APPLICANT: Eveleigh, Deepa
; APPLICANT: Bigwood, Douglas
; TITLE OF INVENTION: EXPRESSION PROFILES FOR BREAST CANCER AND METHODS OF USE
; FILE REFERENCE: 5152
; CURRENT APPLICATION NUMBER: PCT/US04/07268
; PRIOR FILING DATE: 2004-02-27
; PRIOR FILING DATE: 2003-02-28
; NUMBER OF SEQ ID NOS: 254
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 250
; LENGTH: 960
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US04-07268-250

```

late

```

Query Match 99.0%; Score 4983; DB 1; Length 960;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 948; Conservative 0; Mismatches 1; Indels 4; Gaps 4;

Qy 1 MSSGLWSQKVTSPYWEERIFYLLLOBCSVTDKQTKLLKVPKSGIQYIQDRSVGHSRI 60
Db 8 MSSGLWSQKVTSPYWEERIFYLLLOBCSVTDKQTKLLKVPKSGIQYIQDRSVGHSRI 67
Qy 61 PSAKGKKNQIGLKILEQPHAVLFVDE-DVVEINEKEFTTELLAITNCEERFSLFKNRRLS 119
Db 68 PSAKGKKNQIGLKILEQPHAVLFVDEKDVVEINEKEFTTELLAITNCEERFSLFKNRRLS 127
Qy 120 KGLQIDVGCPCVKQLRSGBEEKFPGVVRFRGPLLAERTVSGIFFGVELLEBGRGGFTDGV 179
Db 128 KGLQIDVGCPCVKQLRSGBEEKFPGVVRFRGPLLAERTVSGIFFGVELLEBGRGGFTDGV 187
Qy 180 YGKQLFQDCEDCGVFVLDKLEIEDDDTALSDYAGPDTMQVELPLEINSRVSLKV 238
Db 188 YGKQLFQDCEDCGVFVLDKLEIEDDDTALSDYAGPDTMQVELPLEINSRVSLKV 247
Qy 239 GETIESGTIVFCVLPKESLGYFVGVDMNDPIGNWDRFDGV-LCSFACVESTILLHIN 297
Db 248 GETIESGTIVFCVLPKESLGYFVGVDMNDPIGNWDRFDGV-LCSFACVESTILLHIN 307
Qy 298 DIIPESVTOERRPPKLAFAVRGVDGKSSSHNPKATGSTDGPNR-RSELYFTLNGSSV 356
Db 308 DIIPESVTOERRPPKLAFAVRGVDGKSSSHNPKATGSTDGPNR-RSELYFTLNGSSV 367
Qy 357 DSQPSKSNWTYIDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSLTITENRPHSLPFL 416
Db 368 DSQPSKSNWTYIDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSLTITENRPHSLPFL 427
Qy 417 TKMPNTNGSIGHSPISLSAQSVMEELNAPVQESPPPLAMPNGSHGLEVGSIAEVKENPP 476
Db 428 TKMPNTNGSIGHSPISLSAQSVMEELNAPVQESPPPLAMPNGSHGLEVGSIAEVKENPP 487
Qy 477 FYGVIRWIGQPPGLNEVLAGLEDEBACAGTGTGTRGTRYFTCALKKALFVKLSCRPS 536
Db 488 FYGVIRWIGQPPGLNEVLAGLEDEBACAGTGTGTRGTRYFTCALKKALFVKLSCRPS 547
Qy 537 RFASLOPVSNQIERCNSLAFGGYLSVVEENTPPKMEKEGLEIMIGKKGIQGHYNSCYL 596
Db 548 RFASLOPVSNQIERCNSLAFGGYLSVVEENTPPKMEKEGLEIMIGKKGIQGHYNSCYL 607
Qy 597 DSTLFCFLAFSSVLDTVLLRPKEKNDEYVYSETQELLRTIENVPLRIYGVVCATKIMKLR 656
Db 608 DSTLFCFLAFSSVLDTVLLRPKEKNDEYVYSETQELLRTIENVPLRIYGVVCATKIMKLR 667
Qy 657 KILEKVEAASGFTSEKDPPEEFNLIFHILRVEPLLKIRSAQKQVQDCYFYQIFMEKNE 716
Db 668 KILEKVEAASGFTSEKDPPEEFNLIFHILRVEPLLKIRSAQKQVQDCYFYQIFMEKNE 727
Qy 717 KVGVPITQQLLEWSFINSNLKFAEAPSLIIQMPREGKDFKPKIFPSLELNTDILLED 776
Db 728 KVGVPITQQLLEWSFINSNLKFAEAPSLIIQMPREGKDFKPKIFPSLELNTDILLED 787
Qy 777 TPRQCRICGLAMVYECRECYDDPDISAGKIKQKCTCNTQVHLHPKRLNHNKYNPVSIPKD 836
Db 788 TPRQCRICGLAMVYECRECYDDPDISAGKIKQKCTCNTQVHLHPKRLNHNKYNPVSIPKD 847
Qy 837 LPDMDWRHGCIPCONNELFAVLCTIETSHYVAFVKYKDDSAWLPFDSMADRDGGQGNFI 896
Db 848 LPDMDWRHGCIPCONNELFAVLCTIETSHYVAFVKYKDDSAWLPFDSMADRDGGQGNFI 907
Qy 897 PQVTPCPEVGEYLKMSLEDLHSLDSRRIOGCARRLLCDAYMCMYQSPMTSLYK 949
Db 908 PQVTPCPEVGEYLKMSLEDLHSLDSRRIOGCARRLLCDAYMCMYQSPMTSLYK 960

```

RESULT 6

```

US-10-788-792-250
; Sequence 250, Application US/10788792
; GENERAL INFORMATION:

```


APPLICANT: Bayer Pharmaceuticals Corporation
APPLICANT: Eveleigh, Deepa
APPLICANT: Bigwood, Douglas
TITLE OF INVENTION: EXPRESSION PROFILES FOR BREAST CANCER AND METHODS OF USE
FILE REFERENCE: 5152
CURRENT APPLICATION NUMBER: US/10/788,792
PRIOR FILING DATE: 2004-02-27
PRIOR APPLICATION NUMBER: US 60/450,655
PRIOR FILING DATE: 2003-02-28
NUMBER OF SEQ ID NOS: 254
SOFTWARE: PatentIn version 3.2
SEQ ID NO 250
LENGTH: 960
TYPE: PRT
ORGANISM: Homo sapiens
US-10-788-792-250

Query Match 99.0%; Score 4983; DB 33; Length 960;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 948; Conservative 0; Mismatches 1; Indels 4; Gaps 4;

QY 1 MSSGLWSQEKVTSFYWEERIFVLLQECSTVDKQTKLLKVPKSGSIGQYIQDRSVGHSRI 60
DB 8 MSSGLWSQEKVTSFYWEERIFVLLQECSTVDKQTKLLKVPKSGSIGQYIQDRSVGHSRI 67
QY 61 PSAKKNQIGLKILEQPHAVLFVDE-DVVEINEKFTTELLAITNCEERPSLFKNRNL 119
DB 68 PSAKKNQIGLKILEQPHAVLFVDEKDVVEINEKFTTELLAITNCEERPSLFKNRNL 127
QY 120 KGLQIDVGCPCVKQLRSGEKPGVVRFRGPLLAERTVSGIFPGVELLEGRGQGTG 179
DB 128 KGLQIDVGCPCVKQLRSGEKPGVVRFRGPLLAERTVSGIFPGVELLEGRGQGTG 187
QY 180 YQKQLFQDCEDCG-FVALDKLEIEDDDTALESYAGPGDTMQLPPLPINSRVS 238
DB 188 YQKQLFQDCEDCG-FVALDKLEIEDDDTALESYAGPGDTMQLPPLPINSRVS 247
QY 239 GTTIESGTVIFCDVLPKESLGYFVGVDMDNPIGNWDRFDGV-LCSFACVESTILLH 297
DB 248 GTTIESGTVIFCDVLPKESLGYFVGVDMDNPIGNWDRFDGV-LCSFACVESTILLH 307
QY 298 DIIPESVTOERRPPKLAFAFMRGVDGKSSHNKPKATGSTDPGGR-RSELYFTL 356
DB 308 DIIPESVTOERRPPKLAFAFMRGVDGKSSHNKPKATGSTDPGGR-RSELYFTL 367
QY 357 DSQPSKSNKNTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSLTNENRPHSL 416
DB 368 DSQPSKSNKNTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSLTNENRPHSL 427
QY 417 TKMPNTNGSIGHSPLSLSAQSVMEELNTAPVQESPPPLAMPNGSHGLEVGSLAEV 476
DB 428 TKMPNTNGSIGHSPLSLSAQSVMEELNTAPVQESPPPLAMPNGSHGLEVGSLAEV 487
QY 477 FYGVIRWIOQPGELNEVLAGELEDECACTDGTFRGTRYFTCALKKALFVKL 536
DB 488 FYGVIRWIOQPGELNEVLAGELEDECACTDGTFRGTRYFTCALKKALFVKL 547
QY 537 REASLPQVSNQIERNCSLAFGGVLSWEENTPPKMEKEGLEIMICKKGIQGHYNS 596
DB 548 REASLPQVSNQIERNCSLAFGGVLSWEENTPPKMEKEGLEIMICKKGIQGHYNS 607
QY 597 DSTLFCFLFAFSSVLDVTLRPKEKNVYVSETOELLRTIEINPLRIYGVVCATKIM 656
DB 608 DSTLFCFLFAFSSVLDVTLRPKEKNVYVSETOELLRTIEINPLRIYGVVCATKIM 667
QY 657 KILEKVEAASGFTSEKDPPEFLNLFHILRVEPLLKIRSAQKQVQDCYFYQIFME 716
DB 668 KILEKVEAASGFTSEKDPPEFLNLFHILRVEPLLKIRSAQKQVQDCYFYQIFME 727
QY 717 KVGVPITIOQLLEWFSFNSLNKFAEAPSCLIQMPRFKDFKLPFKIPPSLELNIT 776
DB 728 KVGVPITIOQLLEWFSFNSLNKFAEAPSCLIQMPRFKDFKLPFKIPPSLELNIT 787

777 TPRQCRICGLAMYECECYDDPDIISAGKIQKCTCNTQVHLHPKRLNHNKYNPVS 836
DB 788 TPRQCRICGLAMYECECYDDPDIISAGKIQKCTCNTQVHLHPKRLNHNKYNPVS 847
QY 837 LPDWDWRHGCIPQNMELFAVLCTIETSHYVAFVKYKDDSAWLFDDSMADRDGGQNGFNI 896
DB 848 LPDWDWRHGCIPQNMELFAVLCTIETSHYVAFVKYKDDSAWLFDDSMADRDGGQNGFNI 907
QY 897 PQVTPCEVGEYKMSLEDLHSLDSRRIOGCARRLLCDAYMCYQSTMSLYK 949
DB 908 PQVTPCEVGEYKMSLEDLHSLDSRRIOGCARRLLCDAYMCYQSTMSLYK 960

RESULT 7
PCT-US02-14570-4
Sequence 4, Application PC/TUS0214570
GENERAL INFORMATION:
APPLICANT: IMMUNEX CORPORATION
APPLICANT: Derry, Jonathan M. J.
APPLICANT: Fanslow III, William
APPLICANT: Dougall, William C.
TITLE OF INVENTION: SCREENING ASSAYS FOR AGONISTS OR ANTAGONISTS OF CD40 SIGNALING
FILE REFERENCE: 3198-WO
CURRENT APPLICATION NUMBER: PCT/US02/14570
CURRENT FILING DATE: 2002-05-07
PRIOR APPLICATION NUMBER: 09/851,673
PRIOR FILING DATE: 2001-05-08
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 956
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US02-14570-4

Query Match 98.8%; Score 4971.5; DB 1; Length 956;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 948; Conservative 0; Mismatches 1; Indels 7; Gaps 5;

QY 1 MSSGLWSQEKVTSFYWEERIFVLLQECSTVDKQTKLLKVPKSGSIGQYIQDRSVGHSRI 60
DB 1 MSSGLWSQEKVTSFYWEERIFVLLQECSTVDKQTKLLKVPKSGSIGQYIQDRSVGHSRI 60
QY 61 PSAKKNQIGLKILEQPHAVLFVDE-DVVEINEKFTTELLAITNCEERPSLFKNRNL 119
DB 61 PSAKKNQIGLKILEQPHAVLFVDEKDVVEINEKFTTELLAITNCEERPSLFKNRNL 120
QY 120 KGLQIDVGCPCVKQLRSGEKPGVVRFRGPLLAERTVSGIFPGVELLEGRGQGTG 179
DB 121 KGLQIDVGCPCVKQLRSGEKPGVVRFRGPLLAERTVSGIFPGVELLEGRGQGTG 180
QY 180 YQKQLFQDCEDCG-FVALDKLEIEDDDTALESYAGPGDTMQLPPLPINSRVS 238
DB 181 YQKQLFQDCEDCG-FVALDKLEIEDDDTALESYAGPGDTMQLPPLPINSRVS 240
QY 239 GTTIESGTVIFCDVLPKESLGYFVGVDMDNPIGNWDRFDGV-LCSFACVESTILLH 297
DB 241 GTTIESGTVIFCDVLPKESLGYFVGVDMDNPIGNWDRFDGV-LCSFACVESTILLH 300
QY 298 DIIPESVTOERRPPKLAFAFMRGVDGKSSHNKPKATGSTDPGGR-RSELYFTL 353
DB 301 DIIPESVTOERRPPKLAFAFMRGVDGKSSHNKPKATGSTDPGGR-RSELYFTL 360
QY 354 SSVDSQPSKSNKNTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSLTNENRPHSL 413
DB 361 SSVDSQPSKSNKNTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSLTNENRPHSL 420
QY 414 FSUTKMPNTNGSIGHSPLSLSAQSVMEELNTAPVQESPPPLAMPNGSHGLEVGSLAEV 473
DB 421 FSUTKMPNTNGSIGHSPLSLSAQSVMEELNTAPVQESPPPLAMPNGSHGLEVGSLAEV 480
QY 474 NPPFYGVIRWIOQPGELNEVLAGELEDECACTDGTFRGTRYFTCALKKALFVKL 533

Db 481 NPPFYGVIRWIGQPPGLNEVLAGELEDEACAGCTDGTFRGTRYFTCALKKALFVKLKSCR 540
 QY 534 PDSRFASLOPVNSQIERCNSLAFGGYLSVVEENTPPKMEKEGLEIMIGKKGIQGHYNS 593
 Db 541 PDSRFASLOPVNSQIERCNSLAFGGYLSVVEENTPPKMEKEGLEIMIGKKGIQGHYNS 600
 QY 594 CYLDSTFLCFLAFSSVLDTVLLRPKEKNDVEYYSETQELLRTIENVPLRIYGVVCATKIM 653
 Db 601 CYLDSTFLCFLAFSSVLDTVLLRPKEKNDVEYYSETQELLRTIENVPLRIYGVVCATKIM 660
 QY 654 KLRKILEKVEAASGFTSEEDKPEEFNLILFHILRVEPLLLKRSAGOKVQDCYFYQIFME 713
 Db 661 KLRKILEKVEAASGFTSEEDKPEEFNLILFHILRVEPLLLKRSAGOKVQDCYFYQIFME 720
 QY 714 KNEKVGVPITQQLLEWSFINSNLKFAEAPSCIIQMPRFKGFVKLKKIPPSLELNITDL 773
 Db 721 KNEKVGVPITQQLLEWSFINSNLKFAEAPSCIIQMPRFKGFVKLKKIPPSLELNITDL 780
 QY 774 LEDTPRQCRICGGLAMYECCYDDPDISAGKIKQFCCTCNTQVHLHPKRLNHNKYNPVS 833
 Db 781 LEDTPRQCRICGGLAMYECCYDDPDISAGKIKQFCCTCNTQVHLHPKRLNHNKYNPVS 840
 QY 834 PKDLPDMDWRHGCIPCONMELFAVLCTIETSHYVAFVKYKDDSAWLPFDSMADRGQNG 893
 Db 841 PKDLPDMDWRHGCIPCONMELFAVLCTIETSHYVAFVKYKDDSAWLPFDSMADRGQNG 900
 QY 894 FNIPOVTPCPEVGEYLKMSLEDLHSLDSRRIOGCARRLLCDAYMCYQPTMSLYK 949
 Db 901 FNIPOVTPCPEVGEYLKMSLEDLHSLDSRRIOGCARRLLCDAYMCYQPTMSLYK 956

RESULT 8

US-09-851-673-4

; Sequence 4, Application US/09851673

; GENERAL INFORMATION:

; APPLICANT: Derry, Jonathan

; APPLICANT: Fanslow, William

; APPLICANT: Dougall, William

; TITLE OF INVENTION: SCREENING ASSAYS FOR AGONISTS OR ANTAGONISTS OF CD40 SIGNALING

; FILE REFERENCE: 3198

; CURRENT APPLICATION NUMBER: US/09/851,673

; CURRENT FILING DATE: 2001-05-08

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: Patent in version 3.0

; SEQ ID NO 4

; LENGTH: 956

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-851-673-4

Query Match 98.8%; Score 4971.5; DB 23; Length 956;

Best Local Similarity 99.2%; Pred. No. 0;

Matches 948; Conservative 0; Mismatches 1; Indels 7; Gaps 5;

QY 1 MSSGLWSQEKVTSPPYWEERIFYLLLOECSTVDKOTQKLLKVPKGSIGQYIQDRSVGHSRI 60
 Db 1 MSSGLWSQEKVTSPPYWEERIFYLLLOECSTVDKOTQKLLKVPKGSIGQYIQDRSVGHSRI 60
 QY 61 PSAKGNKQIGLKILQPHAVLPVDE-DVVEINEKTELLLAITNCEERFSLPKNRLS 119
 Db 61 PSAKGNKQIGLKILQPHAVLPVDEKDVVEINEKTELLLAITNCEERFSLPKNRLS 120
 QY 120 KGLQIDVGCVPVKQLSGEKEKFGVVRFCPLAERTVSGIFPGVLLBEGRGQGTGV 179
 Db 121 KGLQIDVGCVPVKQLSGEKEKFGVVRFCPLAERTVSGIFPGVLLBEGRGQGTGV 180
 QY 180 YQKQQLFQDCDCG-FVALDKLEIEDDDTALESYVAGTDMQVLPPLPEINRSVSLKG 238
 Db 181 YQKQQLFQDCDCG-FVALDKLEIEDDDTALESYVAGTDMQVLPPLPEINRSVSLKV 240
 QY 239 GETIESGTVIFCDVLPKESLSGVFVGVMDNPNIGNMDGRPDGV-LCSFACVESTILLHN 297
 Db 241 GETIESGTVIFCDVLPKESLSGVFVGVMDNPNIGNMDGRPDGV-LCSFACVESTILLHN 300

QY 298 DIIP---ESVTOERRPPKLAFMRSRGVGDKGSSSHNKPATGSTDGPNR-RSELFYTLNG 353
 Db 301 DIIPALSESVTQERRPPKLAFMRSRGVGDKGSSSHNKPATGSTDGPNRNRSELFYTLNG 360
 QY 354 SSVDSPQSKSNKWTYIDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSLTITNRRFHSIP 413
 Db 361 SSVDSPQSKSNKWTYIDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSLTITNRRFHSIP 420
 QY 414 FSLTKMPNTNGSIHSPSLSAQSVMEELNTAPQESPPPLAMPNGSHGLEVSLEAVKE 473
 Db 421 FSLTKMPNTNGSIHSPSLSAQSVMEELNTAPQESPPPLAMPNGSHGLEVSLEAVKE 480
 QY 474 NPPFYGVIRWIGQPPGLNEVLAGELEDEACAGCTDGTFRGTRYFTCALKKALFVKLKSCR 533
 Db 481 NPPFYGVIRWIGQPPGLNEVLAGELEDEACAGCTDGTFRGTRYFTCALKKALFVKLKSCR 540
 QY 534 PDSRFASLOPVNSQIERCNSLAFGGYLSVVEENTPPKMEKEGLEIMIGKKGIQGHYNS 593
 Db 541 PDSRFASLOPVNSQIERCNSLAFGGYLSVVEENTPPKMEKEGLEIMIGKKGIQGHYNS 600
 QY 594 CYLDSTFLCFLAFSSVLDTVLLRPKEKNDVEYYSETQELLRTIENVPLRIYGVVCATKIM 653
 Db 601 CYLDSTFLCFLAFSSVLDTVLLRPKEKNDVEYYSETQELLRTIENVPLRIYGVVCATKIM 660
 QY 654 KLRKILEKVEAASGFTSEEDKPEEFNLILFHILRVEPLLLKRSAGOKVQDCYFYQIFME 713
 Db 661 KLRKILEKVEAASGFTSEEDKPEEFNLILFHILRVEPLLLKRSAGOKVQDCYFYQIFME 720
 QY 714 KNEKVGVPITQQLLEWSFINSNLKFAEAPSCIIQMPRFKGFVKLKKIPPSLELNITDL 773
 Db 721 KNEKVGVPITQQLLEWSFINSNLKFAEAPSCIIQMPRFKGFVKLKKIPPSLELNITDL 780
 QY 774 LEDTPRQCRICGGLAMYECCYDDPDISAGKIKQFCCTCNTQVHLHPKRLNHNKYNPVS 833
 Db 781 LEDTPRQCRICGGLAMYECCYDDPDISAGKIKQFCCTCNTQVHLHPKRLNHNKYNPVS 840
 QY 834 PKDLPDMDWRHGCIPCONMELFAVLCTIETSHYVAFVKYKDDSAWLPFDSMADRGQNG 893
 Db 841 PKDLPDMDWRHGCIPCONMELFAVLCTIETSHYVAFVKYKDDSAWLPFDSMADRGQNG 900
 QY 894 FNIPOVTPCPEVGEYLKMSLEDLHSLDSRRIOGCARRLLCDAYMCYQPTMSLYK 949
 Db 901 FNIPOVTPCPEVGEYLKMSLEDLHSLDSRRIOGCARRLLCDAYMCYQPTMSLYK 956

RESULT 9

US-10-755-889-490

; Sequence 490, Application US/10755889

; GENERAL INFORMATION:

; APPLICANT: Bristol-Myers Squibb Company

; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-kB

; FILE REFERENCE: D0284 NP

; CURRENT APPLICATION NUMBER: US/10/755,889

; CURRENT FILING DATE: 2004-01-13

; PRIOR APPLICATION NUMBER: U.S. 60/440,068

; PRIOR FILING DATE: 2003-01-14

; PRIOR APPLICATION NUMBER: U.S. 60/469,757

; PRIOR FILING DATE: 2003-05-12

; NUMBER OF SEQ ID NOS: 823

; SOFTWARE: Patent in version 3.2

; SEQ ID NO 490

; LENGTH: 956

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-755-889-490

Query Match 98.8%; Score 4971.5; DB 33; Length 956;

Best Local Similarity 99.2%; Pred. No. 0;

Matches 948; Conservative 0; Mismatches 1; Indels 7; Gaps 5;

QY 1 MSSGLWSQEKVTSPPYWEERIFYLLLOECSTVDKOTQKLLKVPKGSIGQYIQDRSVGHSRI 60

```
Db 1 MSSGLWSQEKVTSFYWEERIFYLLOQCSVTDKQTKLLKVPKSGISGYIQDRSVGHSRI 60
QY 61 PSAKGKKNQIGLKILEOPHAFLVDE-DVVEINEKFTTELLAITNCEERFSLFKNRRLS 119
Db 61 PSAKGKKNQIGLKILEOPHAFLVDEKDVVEINEKFTTELLAITNCEERFSLFKNRRLS 120
QY 120 KGLQIDVGCPCVKQLASGEEKPGVVRFRGPLLAERTVSGIFPGVLLLEGRGQGTGTV 179
Db 121 KGLQIDVGCPCVKQLASGEEKPGVVRFRGPLLAERTVSGIFPGVLLLEGRGQGTGTV 180
QY 180 YQKQLFQCDDECG-FVALDKLELIEDDDTALSDYAGPDMQVELPPLPINSRVS LKG 238
Db 181 YQKQLFQCDDECGFVALDKLELIEDDDTALSDYAGPDMQVELPPLPINSRVS LKV 240
QY 239 GETIESGTVIFCDVLPKESLGYFVGVDMDNPIGNWDRFDGV-LCSFACVESTILLHIN 297
Db 241 GETIESGTVIFCDVLPKESLGYFVGVDMDNPIGNWDRFDGV-LCSFACVESTILLHIN 300
QY 298 DIIP---ESVTOERRPPKLAFAFMRGVDKSGSSHNKPKATGSTSDPGNR-RSELFYTLNG 353
Db 301 DIIPALSESVTOERRPPKLAFAFMRGVDKSGSSHNKPKATGSTSDPGNRNRS ELFYTLNG 360
QY 354 SSVDSOPQSKSNTWTYIDVAEDPAKSLTEISTDFRSSPPLQPPVNSLTENRPHSLP 413
Db 361 SSVDSOPQSKSNTWTYIDVAEDPAKSLTEISTDFRSSPPLQPPVNSLTENRPHSLP 420
QY 414 FSLTKMPTNGSIGHSPLSLSAQSVMEELNTPVQESPPLAMPNGSHGLEVGS LAEVE 473
Db 421 FSLTKMPTNGSIGHSPLSLSAQSVMEELNTPVQESPPLAMPNGSHGLEVGS LAEVE 480
QY 474 NPPFYGVIRWIGOPGLNEVLAGELEDECACTDGTFRGTRFTYTCALKKALFVKL KSCR 533
Db 481 NPPFYGVIRWIGOPGLNEVLAGELEDECACTDGTFRGTRFTYTCALKKALFVKL KSCR 540
QY 534 PDSRFASLQPVNSQIERCNSLAFGGYLSVEVEENTPPKMEKEGLEIMIGKKKGIQGHYNS 593
Db 541 PDSRFASLQPVNSQIERCNSLAFGGYLSVEVEENTPPKMEKEGLEIMIGKKKGIQGHYNS 600
QY 594 CYLDSTLFCFLAFSSVLDTVLLRPKEKNDEYVYSETQELLRTEIVNPLRIYGVVCATKIM 653
Db 601 CYLDSTLFCFLAFSSVLDTVLLRPKEKNDEYVYSETQELLRTEIVNPLRIYGVVCATKIM 660
QY 654 KLRKILEKVEAASGFTSEKDPPEEFNLIFPHILRVEPLLKIRSAQKQVDCYFYQI FME 713
Db 661 KLRKILEKVEAASGFTSEKDPPEEFNLIFPHILRVEPLLKIRSAQKQVDCYFYQI FME 720
QY 714 KNEKVGVPITIQOLLEWSFINSNLKFAEAPSLIIQMPRFGKDFKPKIFPSLELNI TDL 773
Db 721 KNEKVGVPITIQOLLEWSFINSNLKFAEAPSLIIQMPRFGKDFKPKIFPSLELNI TDL 780
QY 774 LEDTPRQCRI CGGLAMVECRECYDDPDI SAGKIQKQCKTQNTQVHLHPKRLNHNKYNPVS L 833
Db 781 LEDTPRQCRI CGGLAMVECRECYDDPDI SAGKIQKQCKTQNTQVHLHPKRLNHNKYNPVS L 840
QY 834 PKDLPDMDWRHGCIPCONMELFAVLCIETSHYVAFVKYKQKDSAWLFFDSDMADRGQNG 893
Db 841 PKDLPDMDWRHGCIPCONMELFAVLCIETSHYVAFVKYKQKDSAWLFFDSDMADRGQNG 900
QY 894 FNIPOVTPCPEVGEYLKMSLEDLHSLDSRRIOGCARRLLCDAYMCYQSTWSLYK 949
Db 901 FNIPOVTPCPEVGEYLKMSLEDLHSLDSRRIOGCARRLLCDAYMCYQSTWSLYK 956
```

RESULT 10

US-60-440-068-490

```
; Sequence 490, Application US/60440068
; GENERAL INFORMATION:
; APPLICANT: NADLER, STEVEN G.
; APPLICANT: CARMAN, JULIE
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE
; TITLE OF INVENTION: NF-KB PATHWAY
; FILE REFERENCE: 3053-4191
```

```
; CURRENT APPLICATION NUMBER: US/60/440,068
; CURRENT FILING DATE: 2003-01-14
```

; NUMBER OF SEQ ID NOS: 746

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 490

; LENGTH: 956

; TYPE: PRT

; ORGANISM: Homo sapiens

US-60-440-068-490

```
Query Match      98.8%; Score 4971.5; DB 37; Length 956;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 948; Conservative 0; Mismatches 1; Indels 7; Gaps 5;
```

```
QY 1 MSSGLWSQEKVTSFYWEERIFYLLOQCSVTDKQTKLLKVPKSGISGYIQDRSVGHSRI 60
Db 1 MSSGLWSQEKVTSFYWEERIFYLLOQCSVTDKQTKLLKVPKSGISGYIQDRSVGHSRI 60
QY 61 PSAKGKKNQIGLKILEOPHAFLVDE-DVVEINEKFTTELLAITNCEERFSLFKNRRLS 119
Db 61 PSAKGKKNQIGLKILEOPHAFLVDEKDVVEINEKFTTELLAITNCEERFSLFKNRRLS 120
QY 120 KGLQIDVGCPCVKQLASGEEKPGVVRFRGPLLAERTVSGIFPGVLLLEGRGQGTGTV 179
Db 121 KGLQIDVGCPCVKQLASGEEKPGVVRFRGPLLAERTVSGIFPGVLLLEGRGQGTGTV 180
QY 180 YQKQLFQCDDECG-FVALDKLELIEDDDTALSDYAGPDMQVELPPLPINSRVS LKG 238
Db 181 YQKQLFQCDDECGFVALDKLELIEDDDTALSDYAGPDMQVELPPLPINSRVS LKV 240
QY 239 GETIESGTVIFCDVLPKESLGYFVGVDMDNPIGNWDRFDGV-LCSFACVESTILLHIN 297
Db 241 GETIESGTVIFCDVLPKESLGYFVGVDMDNPIGNWDRFDGV-LCSFACVESTILLHIN 300
QY 298 DIIP---ESVTOERRPPKLAFAFMRGVDKSGSSHNKPKATGSTSDPGNR-RSELFYTLNG 353
Db 301 DIIPALSESVTOERRPPKLAFAFMRGVDKSGSSHNKPKATGSTSDPGNRNRS ELFYTLNG 360
QY 354 SSVDSOPQSKSNTWTYIDVAEDPAKSLTEISTDFRSSPPLQPPVNSLTENRPHSLP 413
Db 361 SSVDSOPQSKSNTWTYIDVAEDPAKSLTEISTDFRSSPPLQPPVNSLTENRPHSLP 420
QY 414 FSLTKMPTNGSIGHSPLSLSAQSVMEELNTPVQESPPLAMPNGSHGLEVGS LAEVE 473
Db 421 FSLTKMPTNGSIGHSPLSLSAQSVMEELNTPVQESPPLAMPNGSHGLEVGS LAEVE 480
QY 474 NPPFYGVIRWIGOPGLNEVLAGELEDECACTDGTFRGTRFTYTCALKKALFVKL KSCR 533
Db 481 NPPFYGVIRWIGOPGLNEVLAGELEDECACTDGTFRGTRFTYTCALKKALFVKL KSCR 540
QY 534 PDSRFASLQPVNSQIERCNSLAFGGYLSVEVEENTPPKMEKEGLEIMIGKKKGIQGHYNS 593
Db 541 PDSRFASLQPVNSQIERCNSLAFGGYLSVEVEENTPPKMEKEGLEIMIGKKKGIQGHYNS 600
QY 594 CYLDSTLFCFLAFSSVLDTVLLRPKEKNDEYVYSETQELLRTEIVNPLRIYGVVCATKIM 653
Db 601 CYLDSTLFCFLAFSSVLDTVLLRPKEKNDEYVYSETQELLRTEIVNPLRIYGVVCATKIM 660
QY 654 KLRKILEKVEAASGFTSEKDPPEEFNLIFPHILRVEPLLKIRSAQKQVDCYFYQI FME 713
Db 661 KLRKILEKVEAASGFTSEKDPPEEFNLIFPHILRVEPLLKIRSAQKQVDCYFYQI FME 720
QY 714 KNEKVGVPITIQOLLEWSFINSNLKFAEAPSLIIQMPRFGKDFKPKIFPSLELNI TDL 773
Db 721 KNEKVGVPITIQOLLEWSFINSNLKFAEAPSLIIQMPRFGKDFKPKIFPSLELNI TDL 780
QY 774 LEDTPRQCRI CGGLAMVECRECYDDPDI SAGKIQKQCKTQNTQVHLHPKRLNHNKYNPVS L 833
Db 781 LEDTPRQCRI CGGLAMVECRECYDDPDI SAGKIQKQCKTQNTQVHLHPKRLNHNKYNPVS L 840
QY 834 PKDLPDMDWRHGCIPCONMELFAVLCIETSHYVAFVKYKQKDSAWLFFDSDMADRGQNG 893
Db 841 PKDLPDMDWRHGCIPCONMELFAVLCIETSHYVAFVKYKQKDSAWLFFDSDMADRGQNG 900
```

QY 894 FNIPOVTPCPEVGEYKMSLEDLHSLDSRRIOGCARRLLCDAYMCYQSTMSLYK 949
 Db 901 FNIPOVTPCPEVGEYKMSLEDLHSLDSRRIOGCARRLLCDAYMCYQSTMSLYK 956

RESULT 11
 US-60-469-757-490
 ; Sequence 490, Application US/60469757
 ; GENERAL INFORMATION:
 ; APPLICANT: Bristol-Myers Squibb Company
 ; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-KB
 ; FILE REFERENCE: D0284 PSP1
 ; CURRENT APPLICATION NUMBER: US/60/469,757
 ; NUMBER OF SEQ ID NOS: 823
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 490
 ; LENGTH: 956
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-60-469-757-490

Query Match 98.8%; Score 4971.5; DB 37; Length 956;
 Best Local Similarity 99.2%; Pred. No. 0;
 Matches 948; Conservative 0; Mismatches 1; Indels 7; Gaps 5;

QY 1 MSSGLWSQKVTSPYWEERIFVLLLOECSTVDKOTOKLLKVPKSGISQYIQDRSVGHSRI 60
 Db 1 MSSGLWSQKVTSPYWEERIFVLLLOECSTVDKOTOKLLKVPKSGISQYIQDRSVGHSRI 60

QY 61 PSAKGKNOIGLKILOPHAVLFVDS-DVVEINEKFTELLAITNCEERFSLFKNNRNL 119
 Db 61 PSAKGKNOIGLKILOPHAVLFVDS-DVVEINEKFTELLAITNCEERFSLFKNNRNL 120

QY 120 KGLQIDVGCVPVKVQLRSGBEKPFGVVRFRGPLLAERTVSGIPFGVLLBEGRGQGTG 179
 Db 121 KGLQIDVGCVPVKVQLRSGBEKPFGVVRFRGPLLAERTVSGIPFGVLLBEGRGQGTG 180

QY 180 YQKQLFOCDEDCG-FVALDKLEIEDDDTALESYAGPGDTMQVLPPLPINSRVSLKG 238
 Db 181 YQKQLFOCDEDCG-FVALDKLEIEDDDTALESYAGPGDTMQVLPPLPINSRVSLKV 240

QY 239 GETIESGTVIFCDVLPKESLGVFVGMNDNPIGNWDGRFDGV-LCSFACVESTILLHN 297
 Db 241 GETIESGTVIFCDVLPKESLGVFVGMNDNPIGNWDGRFDGV-LCSFACVESTILLHN 300

QY 298 DIIP-----ESVTQERRPPKLAFMRSRGVDKSSSHNPKATGSTSDPGNR-RSELFYTLNG 353
 Db 301 DIIPALSESVTQERRPPKLAFMRSRGVDKSSSHNPKATGSTSDPGNRNRSFLYTLNG 360

QY 354 SSVDSQPSQSKNTWIDEVAEDPAKSLTEISTDFRSPPLPQPPVNSLTNRRPHSLP 413
 Db 361 SSVDSQPSQSKNTWIDEVAEDPAKSLTEISTDFRSPPLPQPPVNSLTNRRPHSLP 420

QY 414 FSLTKMNTNGSCHSPSLSAQSVMEELNTPVQESPPPLAMPNGSHGLEVGSIAEYKE 473
 Db 421 FSLTKMNTNGSCHSPSLSAQSVMEELNTPVQESPPPLAMPNGSHGLEVGSIAEYKE 480

QY 474 NPPFYGVIRWIGPPGLNEVLAGELEDEACAGCTDGTFRGTRYFTCALKKALFVKLKSCR 533
 Db 481 NPPFYGVIRWIGPPGLNEVLAGELEDEACAGCTDGTFRGTRYFTCALKKALFVKLKSCR 540

QY 534 PDSRFASLPVSNQIERCNSLAFGGYLSVVEENTPPKMEKEGLEIMIGKKGIQGHYNS 593
 Db 541 PDSRFASLPVSNQIERCNSLAFGGYLSVVEENTPPKMEKEGLEIMIGKKGIQGHYNS 600

QY 594 CYLDSTLFCFLFAFSSVLDTVLLRPKKNQVVEYSETQELLRTETVNPRLRYGVVCATKIM 653
 Db 601 CYLDSTLFCFLFAFSSVLDTVLLRPKKNQVVEYSETQELLRTETVNPRLRYGVVCATKIM 660

QY 654 KLRKILEKVEAASGFTSEBKDPPEFLNLPFHLLRVEPLPKIRSAQKVQDCFYQIFME 713

Db 661 KURKILEKVEAASGFTSEBKDPPEFLNLPFHLLRVEPLPKIRSAQKVQDCFYQIFME 720
 QY 714 KNEKVGVPITIQOLLEWSFINSLKFAEAPSCSLIIQMPRFGKDFKLFKKIIPFSLNITDL 773
 Db 721 KNEKVGVPITIQOLLEWSFINSLKFAEAPSCSLIIQMPRFGKDFKLFKKIIPFSLNITDL 780

QY 774 LEDTPRQCRICGGLAMYECCREYDDPDISAGKIKQFCCKTQNTQVHLHPKLNHNKYNPVS 833
 Db 781 LEDTPRQCRICGGLAMYECCREYDDPDISAGKIKQFCCKTQNTQVHLHPKLNHNKYNPVS 840

QY 834 PKDLDPMDWRHGCIPCONMELFAVLCTETSHYVAFVKYKGDSDAWLFFDSMARDGGONG 893
 Db 841 PKDLDPMDWRHGCIPCONMELFAVLCTETSHYVAFVKYKGDSDAWLFFDSMARDGGONG 900

QY 894 FNIPOVTPCPEVGEYKMSLEDLHSLDSRRIOGCARRLLCDAYMCYQSTMSLYK 949
 Db 901 FNIPOVTPCPEVGEYKMSLEDLHSLDSRRIOGCARRLLCDAYMCYQSTMSLYK 956

RESULT 12
 PCT-US01-01239-1743
 ; Sequence 1743, Application PC/TUS0101239
 ; GENERAL INFORMATION:
 ; APPLICANT: Human Genome Sciences, Inc., et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PT213PCT
 ; CURRENT APPLICATION NUMBER: PCT/US01/01239
 ; CURRENT FILING DATE: 2001-01-17
 ; Prior application data removed - refer to PALM or file wrapper
 ; NUMBER OF SEQ ID NOS: 2318
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 1743
 ; LENGTH: 739
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 PCT-US01-01239-1743

Query Match 77.0%; Score 3876.5; DB 1; Length 739;
 Best Local Similarity 95.7%; Pred. No. 0;
 Matches 736; Conservative 0; Mismatches 0; Indels 33; Gaps 4;

QY 184 QLFQCDDEDCG-FVALDKLEIEDDDTALESYAGPGDTMQVLPPLPINSRVSLKGETI 242
 Db 1 QLFQCDDEDCG-FVALDKLEIEDDDTALESYAGPGDTMQVLPPLPINSRVSLK- 55

QY 243 ESGTVIFCDVLPKESLGVFVGMNDNPIGNWDGRFDGV-LCSFACVESTILLHINDIIP 301
 Db 56 -----DNPIGNWDGRFDGVQLCSFACVESTILLHINDIIP 90

QY 302 ESVTQERRPPKLAFMRSRGVDKSSSHNPKATGSTSDPGNR-RSELFYTLNGSSVDSQP 360
 Db 91 ESVTQERRPPKLAFMRSRGVDKSSSHNPKATGSTSDPGNRNRSFLYTLNGSSVDSQP 150

QY 361 QSKSKNTWIDEVAEDPAKSLTEISTDFRSPPLPQPPVNSLTNRRPHSLPFLTKMP 420
 Db 151 QSKSKNTWIDEVAEDPAKSLTEISTDFRSPPLPQPPVNSLTNRRPHSLPFLTKMP 210

QY 421 NTNGSIGHSPSLSLSAQSVMEELNTPVQESPPPLAMPNGSHGLEVGSIAEVENPPFYGV 480
 Db 211 NTNGSIGHSPSLSLSAQSVMEELNTPVQESPPPLAMPNGSHGLEVGSIAEVENPPFYGV 270

QY 481 IRWIGOPPLNEVLAGELEDEACAGCTDGTFRGTRYFTCALKKALFVKLKSCRDPSPAS 540
 Db 271 IRWIGOPPLNEVLAGELEDEACAGCTDGTFRGTRYFTCALKKALFVKLKSCRDPSPAS 330

QY 541 LQPVSNQIERCNSLAFGGYLSVVEENTPPKMEKEGLEIMIGKKGIQGHYNSCYLDSTL 600
 Db 331 LQPVSNQIERCNSLAFGGYLSVVEENTPPKMEKEGLEIMIGKKGIQGHYNSCYLDSTL 390

QY 601 FCLFAFSSVLDTVLLRPKKNQVVEYSETQELLRTETVNPRLRYGVVCATKIMKRLKILE 660
 Db 391 FCLFAFSSVLDTVLLRPKKNQVVEYSETQELLRTETVNPRLRYGVVCATKIMKRLKILE 450

QY 661 KVEAASGFTSEEDKPEEFNLHILRVBPLKIRSAQKQVDCYFYQIFMEKNEKGV 720
DB 451 KVEAASGFTSEEDKPEEFNLHILRVBPLKIRSAQKQVDCYFYQIFMEKNEKGV 510
QY 721 PTIQOLLEWSFINSNLKFAEAPSCILIIQMPRFGKDFKLPKIPPSLELNITDLEDTPRQ 780
DB 511 PTIQOLLEWSFINSNLKFAEAPSCILIIQMPRFGKDFKLPKIPPSLELNITDLEDTPRQ 570
QY 781 CRICGGLAMEYECRECYDDPDISAGKIKQFCKTCTQVHLHPKRLNHNKYNPVSPLPKDLPDW 840
DB 571 CRICGGLAMEYECRECYDDPDISAGKIKQFCKTCTQVHLHPKRLNHNKYNPVSPLPKDLPDW 630
QY 841 DWRHGCIPQCNMELFAVLCTIETSHYVAFVYKGDSDSAWLPFDSMADRDGGQNGFNIPQVT 900
DB 631 DWRHGCIPQCNMELFAVLCTIETSHYVAFVYKGDSDSAWLPFDSMADRDGGQNGFNIPQVT 690
QY 901 PCPEVGEYLKMSLEDLHSLDSRRIOGCARRLLCDAYMCYQSPMTSLYK 949
DB 691 PCPEVGEYLKMSLEDLHSLDSRRIOGCARRLLCDAYMCYQSPMTSLYK 739
RESULT 13
US-09-764-902-1743
; Sequence 1743, Application US/09764902
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT213
; CURRENT APPLICATION NUMBER: US/09/764,902
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2318
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1743
; LENGTH: 739
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-902-1743
Query Match 77.0%; Score 3876.5; DB 22; Length 739;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 736; Conservative 0; Mismatches 0; Indels 33; Gaps 4;

QY 184 QLFQCDDECG-FVALDKLEIEDDDTALSDYAGPGDTMQVELPPLPINSVSLKGGETI 242
DB 1 QLFQCDDECGFVALDKLEIEDDDTALSDYAGPGDTMQVELPPLPINSVSLK---- 55
QY 243 ESGTVIFCDVLPKESLGYFVGVDMDNPIGNWDGRFDGV-LCSFACVESTILLHINDIIP 301
DB 56 -----DNPICGNWDGRFDGVQLCSFACVESTILLHINDIIP 90
QY 302 ESVTOERRPPKLAFMRSRGVGDGSSSHNPKATGSTSDPCNR-RSELFTLANGSSVDSOP 360
DB 91 ESVTOERRPPKLAFMRSRGVGDGSSSHNPKATGSTSDPCNRRSELFTLANGSSVDSOP 150
QY 361 QSKSKNTWYIDEVAEDPAKSLTEISDFDRSGPPLQPPVNSLTTRNPHSLPFSITKTP 420
DB 151 QSKSKNTWYIDEVAEDPAKSLTEISDFDRSGPPLQPPVNSLTTRNPHSLPFSITKTP 210
QY 421 NTNGSIGHSPLSLSAQSVMEELNTPVQSPPLAMPNGSHGLEVGSIAEVKENPPFYGV 480
DB 211 NTNGSIGHSPLSLSAQSVMEELNTPVQSPPLAMPNGSHGLEVGSIAEVKENPPFYGV 270
QY 481 IRWICQPPGLNEVLAGELEDEACAGTDTGTFRGTRYFTTALKKALPVKLSKCRPDRFPAS 540
DB 271 IRWICQPPGLNEVLAGELEDEACAGTDTGTFRGTRYFTTALKKALPVKLSKCRPDRFPAS 330
QY 541 LQPVSNQIERCNLSAFGGVLSVVEENTPPKMEKEGLEIMIGKKGIQGHYNSCYLDSTL 600
DB 331 LQPVSNQIERCNLSAFGGVLSVVEENTPPKMEKEGLEIMIGKKGIQGHYNSCYLDSTL 390
QY 601 FCLFAFSSVLDTVLLRPKEKNDVEYYSETQELLRTBIIVNPLRIYGVVCAIKMLRKILE 660

DB 391 FCLFAFSSVLDTVLLRPKEKNDVEYYSETQELLRTBIIVNPLRIYGVVCAIKMLRKILE 450
QY 661 KVEAASGFTSEEDKPEEFNLHILRVBPLKIRSAQKQVDCYFYQIFMEKNEKGV 720
DB 451 KVEAASGFTSEEDKPEEFNLHILRVBPLKIRSAQKQVDCYFYQIFMEKNEKGV 510
QY 721 PTIQOLLEWSFINSNLKFAEAPSCILIIQMPRFGKDFKLPKIPPSLELNITDLEDTPRQ 780
DB 511 PTIQOLLEWSFINSNLKFAEAPSCILIIQMPRFGKDFKLPKIPPSLELNITDLEDTPRQ 570
QY 781 CRICGGLAMEYECRECYDDPDISAGKIKQFCKTCTQVHLHPKRLNHNKYNPVSPLPKDLPDW 840
DB 571 CRICGGLAMEYECRECYDDPDISAGKIKQFCKTCTQVHLHPKRLNHNKYNPVSPLPKDLPDW 630
QY 841 DWRHGCIPQCNMELFAVLCTIETSHYVAFVYKGDSDSAWLPFDSMADRDGGQNGFNIPQVT 900
DB 631 DWRHGCIPQCNMELFAVLCTIETSHYVAFVYKGDSDSAWLPFDSMADRDGGQNGFNIPQVT 690
QY 901 PCPEVGEYLKMSLEDLHSLDSRRIOGCARRLLCDAYMCYQSPMTSLYK 949
DB 691 PCPEVGEYLKMSLEDLHSLDSRRIOGCARRLLCDAYMCYQSPMTSLYK 739
RESULT 14
US-09-786-797B-9
; Sequence 9, Application US/09786797B
; GENERAL INFORMATION:
; APPLICANT: INCYTE PHARMACEUTICALS, INC.
; APPLICANT: LAL, Preeti
; APPLICANT: TANG, Y. Tom
; APPLICANT: YUE, Henry
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: BANDMAN, Olga
; APPLICANT: CORLEY, Neil C.
; APPLICANT: GUEGLER, Karl J.
; APPLICANT: PATTERSON, Chandra
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: BAUGHN, Mariah R.
; TITLE OF INVENTION: HUMAN CYTOSKELETON ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0594 PCT
; CURRENT APPLICATION NUMBER: US/09/786,797B
; CURRENT FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: 09/156,470; unassigned; 60/131,321
; PRIOR FILING DATE: 1998-09-18; 1998-09-18; 1999-04-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PERL Program
; SEQ ID NO 9
; LENGTH: 731
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 2363327
US-09-786-797B-9
Query Match 76.7%; Score 3862; DB 22; Length 731;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 728; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY 221 MQVELPPLPINSRVSLKGGETIESGTIVFCDLVPGKESLGYFVGVDMDNPIGNWDGRPDG 280
DB 1 MQVELPPLPINSRVSLKGGETIESGTIVFCDLVPGKESLGYFVGVDMDNPIGNWDGRPDG 60
QY 281 V-LCSFACVESTILLHINDIIPESVTOERRPPKLAFMRSRGVGDGSSSHNPKATGSTSD 339
DB 61 VQLCSFACVESTILLHINDIIPESVTOERRPPKLAFMRSRGVGDGSSSHNPKATGSTSD 120
QY 340 PGNR-RSELFTLANGSSVDSOPQSKNTWYIDEVAEDPAKSLTEISDFDRSGPPLQPP 398
DB 121 PGNRRSELFTLANGSSVDSOPQSKNTWYIDEVAEDPAKSLTEISDFDRSGPPLQPP 180
QY 399 PVNSLTTRNPHSLPFSITKMPNTNGSIGHSPLSLSAQSVMEELNTPVQSPPLAMP 458

```
Db 181 PVSNTTNRHSLPFLSLTKMNTNGSIHSPSLSLSAQSYMELNTPAQESPPPLAMPFG 240
Qy 459 NSHGLEVSLAEVKNPPYGVIRWIGQPPGLNEVLAGELEDEACAGCTDGTFRGTRYFT 518
Db 241 NSHGLEVSLAEVKNPPYGVIRWIGQPPGLNEVLAGELEDEACAGCTDGTFRGTRYFT 300
Qy 519 CALKKALFVKLKSRRPDRSFASLQPVSNQIERCNSLAFGGYLSVVVEENTPPKMEKEGLE 578
Db 301 CALKKALFVKLKSRRPDRSFASLQPVSNQIERCNSLAFGGYLSVVVEENTPPKMEKEGLE 360
Qy 579 IMIGKKGIQGHYNSCYLDSTLFLCLFAFSSVLDTVLLRPKEKNDVEYYSETQELLRTETV 638
Db 361 IMIGKKGIQGHYNSCYLDSTLFLCLFAFSSVLDTVLLRPKEKNDVEYYSETQELLRTETV 420
Qy 639 NPLRIYGVYCATKIMKRLKLEKVAASGFTSEEDPPEFLNLFPHILRVEPLLKIRSA 698
Db 421 NPLRIYGVYCATKIMKRLKLEKVAASGFTSEEDPPEFLNLFPHILRVEPLLKIRSA 480
Qy 699 GOKVQDCYFYQIFMEKNEKVGVTIQQLLEWSFINSNLKFAEAPSCLIIQMPRFGKDFKL 758
Db 481 GOKVQDCYFYQIFMEKNEKVGVTIQQLLEWSFINSNLKFAEAPSCLIIQMPRFGKDFKL 540
Qy 759 FKIIFFSLNITDLEDTTPROCRICGGLAMYECECYDDPDISAGKIKQFCKTCTQVH 818
Db 541 FKIIFFSLNITDLEDTTPROCRICGGLAMYECECYDDPDISAGKIKQFCKTCTQVH 600
Qy 819 LHPKRLNHNKYNPVSIPKDLPDMDWRHGCIPQNMELFAVLCTIETSHYVAFVYKGDSDSAW 878
Db 601 LHPKRLNHNKYNPVSIPKDLPDMDWRHGCIPQNMELFAVLCTIETSHYVAFVYKGDSDSAW 660
Qy 879 LFFDSMADRGGQNGFNIPQVTPCPEVGEYVKMSLEDLHSLDSRRIQGCARRLLCDAYMC 938
Db 661 LFFDSMADRGGQNGFNIPQVTPCPEVGEYVKMSLEDLHSLDSRRIQGCARRLLCDAYMC 720
Qy 939 MYQSPMTSLYK 949
Db 721 MYQSPMTSLYK 731
```

RESULT 15

```
US-10-921-707-9
; Sequence 9, Application US/10921707
; GENERAL INFORMATION:
; APPLICANT: INCYTE PHARMACEUTICALS, INC.
; APPLICANT: LAL, Preeti
; APPLICANT: TANG, Y. Tom
; APPLICANT: YUE, Henry
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: BANDMAN, Olga
; APPLICANT: CORLEY, Neil C.
; APPLICANT: GUEGLER, Karl J.
; APPLICANT: PATTERSON, Chandra
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: BAUGHN, Mariah R.
; TITLE OF INVENTION: HUMAN CYTOSKELETON ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0594 PCT
; CURRENT APPLICATION NUMBER: US/10/921,707
; CURRENT FILING DATE: 2004-08-19
; PRIOR APPLICATION NUMBER: US/09/786,797
; PRIOR FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: 09/156,470; unassigned; 60/131,321
; PRIOR FILING DATE: 1998-09-18; 1998-09-18; 1999-04-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PERL Program
; SEQ ID NO 9
; LENGTH: 731
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 2363327
US-10-921-707-9
```

```
Query Match 76.7%; Score 3862; DB 35; Length 731;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 728; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

Qy 221 MQVELPPLINRSVSLKGGTIESGTVIFCDVLPKESLGYPVGVDMNDPIGNWDGDFDG 280
Db 1 MQVELPPLINRSVSLKGGTIESGTVIFCDVLPKESLGYPVGVDMNDPIGNWDGDFDG 60
Qy 281 V-LCSPACVESTILLHNDIIIPESVTQERRPPKLPFMSRGVGDGKSSSHNKPATGSTSD 339
Db 61 VQLCSPACVESTILLHNDIIIPESVTQERRPPKLPFMSRGVGDGKSSSHNKPATGSTSD 120
Qy 340 PGNR-RSELFYTLINGSSVDSQPSQSKNTWYIDEVAEDPAKSLTEISTDFDRSSPPIQPP 398
Db 121 PGNR-RSELFYTLINGSSVDSQPSQSKNTWYIDEVAEDPAKSLTEISTDFDRSSPPIQPP 180
Qy 399 PVSNTTNRHSLPFLSLTKMNTNGSIHSPSLSLSAQSYMELNTPAQESPPPLAMPFG 458
Db 181 PVSNTTNRHSLPFLSLTKMNTNGSIHSPSLSLSAQSYMELNTPAQESPPPLAMPFG 240
Qy 459 NSHGLEVSLAEVKNPPYGVIRWIGQPPGLNEVLAGELEDEACAGCTDGTFRGTRYFT 518
Db 241 NSHGLEVSLAEVKNPPYGVIRWIGQPPGLNEVLAGELEDEACAGCTDGTFRGTRYFT 300
Qy 519 CALKKALFVKLKSRRPDRSFASLQPVSNQIERCNSLAFGGYLSVVVEENTPPKMEKEGLE 578
Db 301 CALKKALFVKLKSRRPDRSFASLQPVSNQIERCNSLAFGGYLSVVVEENTPPKMEKEGLE 360
Qy 579 IMIGKKGIQGHYNSCYLDSTLFLCLFAFSSVLDTVLLRPKEKNDVEYYSETQELLRTETV 638
Db 361 IMIGKKGIQGHYNSCYLDSTLFLCLFAFSSVLDTVLLRPKEKNDVEYYSETQELLRTETV 420
Qy 639 NPLRIYGVYCATKIMKRLKLEKVAASGFTSEEDPPEFLNLFPHILRVEPLLKIRSA 698
Db 421 NPLRIYGVYCATKIMKRLKLEKVAASGFTSEEDPPEFLNLFPHILRVEPLLKIRSA 480
Qy 699 GOKVQDCYFYQIFMEKNEKVGVTIQQLLEWSFINSNLKFAEAPSCLIIQMPRFGKDFKL 758
Db 481 GOKVQDCYFYQIFMEKNEKVGVTIQQLLEWSFINSNLKFAEAPSCLIIQMPRFGKDFKL 540
Qy 759 FKIIFFSLNITDLEDTTPROCRICGGLAMYECECYDDPDISAGKIKQFCKTCTQVH 818
Db 541 FKIIFFSLNITDLEDTTPROCRICGGLAMYECECYDDPDISAGKIKQFCKTCTQVH 600
Qy 819 LHPKRLNHNKYNPVSIPKDLPDMDWRHGCIPQNMELFAVLCTIETSHYVAFVYKGDSDSAW 878
Db 601 LHPKRLNHNKYNPVSIPKDLPDMDWRHGCIPQNMELFAVLCTIETSHYVAFVYKGDSDSAW 660
Qy 879 LFFDSMADRGGQNGFNIPQVTPCPEVGEYVKMSLEDLHSLDSRRIQGCARRLLCDAYMC 938
Db 661 LFFDSMADRGGQNGFNIPQVTPCPEVGEYVKMSLEDLHSLDSRRIQGCARRLLCDAYMC 720
Qy 939 MYQSPMTSLYK 949
Db 721 MYQSPMTSLYK 731
```

Search completed: April 18, 2005, 15:32:41
Job time : 359 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 18, 2005, 15:35:41 ; Search time 7161 Seconds
(without alignments)
5410.008 Million cell updates/sec

Title: US-09-671-687A-3

Perfect score: 5034

Sequence: 1 MSSGLWSQKVTSPWBERI.....RLLCDAYMCYQPTWSLYK 949

Scoring table:

Xgapop 10.0 ,	Xgapext 0.5
Ygapop 10.0 ,	Ygapext 0.5
Fgapop 6.0 ,	Fgapext 7.0
Delop 6.0 ,	Delext 7.0

Searched: 45554873 seqs, 20411521753 residues

Total number of hits satisfying chosen parameters: 91109746

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame-p2n.model -DEV=xlp
-O=/cgn2_1/USPTO_spool_p/US09671687/runat_18042005_115019_17191/app_query.fasta_1.1095
-DB=Pending_Patents_NA_Main_QPMT=fastap -SUFFIX=rnnp -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MAIRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pet -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09671687 @CGN 1 1 6376 @runat_18042005_115019_17191 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: Pending Patents NA_Main:*
2: /cgn2_6/ptodata/1/pna/PCTUS1_COMB.seq:*
3: /cgn2_6/ptodata/1/pna/PCTUS2_COMB.seq:*
4: /cgn2_6/ptodata/1/pna/US06_COMB.seq:*
5: /cgn2_6/ptodata/1/pna/US07_COMB.seq:*
6: /cgn2_6/ptodata/1/pna/US080_COMB.seq:*
7: /cgn2_6/ptodata/1/pna/US081_COMB.seq:*
8: /cgn2_6/ptodata/1/pna/US082_COMB.seq:*
9: /cgn2_6/ptodata/1/pna/US083_COMB.seq:*
10: /cgn2_6/ptodata/1/pna/US084_COMB.seq:*
11: /cgn2_6/ptodata/1/pna/US085_COMB.seq:*
12: /cgn2_6/ptodata/1/pna/US086_COMB.seq:*
13: /cgn2_6/ptodata/1/pna/US087_COMB.seq:*
14: /cgn2_6/ptodata/1/pna/US088_COMB.seq:*
15: /cgn2_6/ptodata/1/pna/US089_COMB.seq:*
16: /cgn2_6/ptodata/1/pna/US090_COMB.seq:*
17: /cgn2_6/ptodata/1/pna/US091_COMB.seq:*
18: /cgn2_6/ptodata/1/pna/US092A_COMB.seq:*
19: /cgn2_6/ptodata/1/pna/US092B_COMB.seq:*
20: /cgn2_6/ptodata/1/pna/US093A_COMB.seq:*
21: /cgn2_6/ptodata/1/pna/US093B_COMB.seq:*
22: /cgn2_6/ptodata/1/pna/US094_COMB.seq:*
23: /cgn2_6/ptodata/1/pna/US095A_COMB.seq:*
24: /cgn2_6/ptodata/1/pna/US095B_COMB.seq:*
25: /cgn2_6/ptodata/1/pna/US095C_COMB.seq:*
26: /cgn2_6/ptodata/1/pna/US095D_COMB.seq:*
27: /cgn2_6/ptodata/1/pna/US096A_COMB.seq:*
28: /cgn2_6/ptodata/1/pna/US096B_COMB.seq:*

29: /cgn2_6/ptodata/1/pna/US096C_COMB.seq:*
30: /cgn2_6/ptodata/1/pna/US096D_COMB.seq:*
31: /cgn2_6/ptodata/1/pna/US096E_COMB.seq:*
32: /cgn2_6/ptodata/1/pna/US097A_COMB.seq:*
33: /cgn2_6/ptodata/1/pna/US097B_COMB.seq:*
34: /cgn2_6/ptodata/1/pna/US097C_COMB.seq:*
35: /cgn2_6/ptodata/1/pna/US098A_COMB.seq:*
36: /cgn2_6/ptodata/1/pna/US098B_COMB.seq:*
37: /cgn2_6/ptodata/1/pna/US098C_COMB.seq:*
38: /cgn2_6/ptodata/1/pna/US098D_COMB.seq:*
39: /cgn2_6/ptodata/1/pna/US099A_COMB.seq:*
40: /cgn2_6/ptodata/1/pna/US099B_COMB.seq:*
41: /cgn2_6/ptodata/1/pna/US099C_COMB.seq:*
42: /cgn2_6/ptodata/1/pna/US099D_COMB.seq:*
43: /cgn2_6/ptodata/1/pna/US099E_COMB.seq:*
44: /cgn2_6/ptodata/1/pna/US099F_COMB.seq:*
45: /cgn2_6/ptodata/1/pna/US099G_COMB.seq:*
46: /cgn2_6/ptodata/1/pna/US100A_COMB.seq:*
47: /cgn2_6/ptodata/1/pna/US100B_COMB.seq:*
48: /cgn2_6/ptodata/1/pna/US100C_COMB.seq:*
49: /cgn2_6/ptodata/1/pna/US101A_COMB.seq:*
50: /cgn2_6/ptodata/1/pna/US101B_COMB.seq:*
51: /cgn2_6/ptodata/1/pna/US102A_COMB.seq:*
52: /cgn2_6/ptodata/1/pna/US102B_COMB.seq:*
53: /cgn2_6/ptodata/1/pna/US103A_COMB.seq:*
54: /cgn2_6/ptodata/1/pna/US104A_COMB.seq:*
55: /cgn2_6/ptodata/1/pna/US104B_COMB.seq:*
56: /cgn2_6/ptodata/1/pna/US105A_COMB.seq:*
57: /cgn2_6/ptodata/1/pna/US105B_COMB.seq:*
58: /cgn2_6/ptodata/1/pna/US106A_COMB.seq:*
59: /cgn2_6/ptodata/1/pna/US107A_COMB.seq:*
60: /cgn2_6/ptodata/1/pna/US107B_COMB.seq:*
61: /cgn2_6/ptodata/1/pna/US107C_COMB.seq:*
62: /cgn2_6/ptodata/1/pna/US107D_COMB.seq:*
63: /cgn2_6/ptodata/1/pna/US108A_COMB.seq:*
64: /cgn2_6/ptodata/1/pna/US108B_COMB.seq:*
65: /cgn2_6/ptodata/1/pna/US109A_COMB.seq:*
66: /cgn2_6/ptodata/1/pna/US109B_COMB.seq:*
67: /cgn2_6/ptodata/1/pna/US109C_COMB.seq:*
68: /cgn2_6/ptodata/1/pna/US110_COMB.seq:*
69: /cgn2_6/ptodata/1/pna/US6000_COMB.seq:*
70: /cgn2_6/ptodata/1/pna/US6001_COMB.seq:*
71: /cgn2_6/ptodata/1/pna/US6002_COMB.seq:*
72: /cgn2_6/ptodata/1/pna/US6003_COMB.seq:*
73: /cgn2_6/ptodata/1/pna/US6004_COMB.seq:*
74: /cgn2_6/ptodata/1/pna/US6005_COMB.seq:*
75: /cgn2_6/ptodata/1/pna/US6006_COMB.seq:*
76: /cgn2_6/ptodata/1/pna/US6007_COMB.seq:*
77: /cgn2_6/ptodata/1/pna/US6008_COMB.seq:*
78: /cgn2_6/ptodata/1/pna/US6009_COMB.seq:*
79: /cgn2_6/ptodata/1/pna/US6010_COMB.seq:*
80: /cgn2_6/ptodata/1/pna/US6011_COMB.seq:*
81: /cgn2_6/ptodata/1/pna/US6012_COMB.seq:*
82: /cgn2_6/ptodata/1/pna/US6013_COMB.seq:*
83: /cgn2_6/ptodata/1/pna/US6014_COMB.seq:*
84: /cgn2_6/ptodata/1/pna/US6015_COMB.seq:*
85: /cgn2_6/ptodata/1/pna/US6016_COMB.seq:*
86: /cgn2_6/ptodata/1/pna/US6017_COMB.seq:*
87: /cgn2_6/ptodata/1/pna/US6018_COMB.seq:*
88: /cgn2_6/ptodata/1/pna/US6019_COMB.seq:*
89: /cgn2_6/ptodata/1/pna/US6020_COMB.seq:*
90: /cgn2_6/ptodata/1/pna/US6021_COMB.seq:*
91: /cgn2_6/ptodata/1/pna/US6022_COMB.seq:*
92: /cgn2_6/ptodata/1/pna/US6023A_COMB.seq:*
93: /cgn2_6/ptodata/1/pna/US6023B_COMB.seq:*
94: /cgn2_6/ptodata/1/pna/US6024_COMB.seq:*
95: /cgn2_6/ptodata/1/pna/US6025_COMB.seq:*
96: /cgn2_6/ptodata/1/pna/US6026_COMB.seq:*
97: /cgn2_6/ptodata/1/pna/US6027_COMB.seq:*
98: /cgn2_6/ptodata/1/pna/US6028_COMB.seq:*
99: /cgn2_6/ptodata/1/pna/US6029_COMB.seq:*
100: /cgn2_6/ptodata/1/pna/US6030_COMB.seq:*
101: /cgn2_6/ptodata/1/pna/US6031_COMB.seq:*

102: /cgn2_6/ptodata/1/pna/US6032_COMB.seq.*
 103: /cgn2_6/ptodata/1/pna/US6033_COMB.seq.*
 104: /cgn2_6/ptodata/1/pna/US6034_COMB.seq.*
 105: /cgn2_6/ptodata/1/pna/US6035_COMB.seq.*
 106: /cgn2_6/ptodata/1/pna/US6036_COMB.seq.*
 107: /cgn2_6/ptodata/1/pna/US6037_COMB.seq.*
 108: /cgn2_6/ptodata/1/pna/US6038_COMB.seq.*
 109: /cgn2_6/ptodata/1/pna/US6039_COMB.seq.*
 110: /cgn2_6/ptodata/1/pna/US6040_COMB.seq.*
 111: /cgn2_6/ptodata/1/pna/US6041_COMB.seq.*
 112: /cgn2_6/ptodata/1/pna/US6042_COMB.seq.*
 113: /cgn2_6/ptodata/1/pna/US6043_COMB.seq.*
 114: /cgn2_6/ptodata/1/pna/US6044_COMB.seq.*
 115: /cgn2_6/ptodata/1/pna/US6045_COMB.seq.*
 116: /cgn2_6/ptodata/1/pna/US6046_COMB.seq.*
 117: /cgn2_6/ptodata/1/pna/US6047_COMB.seq.*
 118: /cgn2_6/ptodata/1/pna/US6048_COMB.seq.*
 119: /cgn2_6/ptodata/1/pna/US6049_COMB.seq.*
 120: /cgn2_6/ptodata/1/pna/US6050_COMB.seq.*
 121: /cgn2_6/ptodata/1/pna/US6051_COMB.seq.*
 122: /cgn2_6/ptodata/1/pna/US6052_COMB.seq.*
 123: /cgn2_6/ptodata/1/pna/US6053_COMB.seq.*
 124: /cgn2_6/ptodata/1/pna/US6054_COMB.seq.*
 125: /cgn2_6/ptodata/1/pna/US6055_COMB.seq.*
 126: /cgn2_6/ptodata/1/pna/US6056_COMB.seq.*
 127: /cgn2_6/ptodata/1/pna/US6057_COMB.seq.*
 128: /cgn2_6/ptodata/1/pna/US6058_COMB.seq.*
 129: /cgn2_6/ptodata/1/pna/US6059_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	4992	99.2	3715	30	US-09-671-687A-2
2	4983	99.0	4661	23	US-09-522-303-1293
3	4983	99.0	4664	20	US-09-315-788-1729
4	4983	99.0	4664	20	US-09-315-788-1729
5	4983	99.0	4664	34	US-09-783-514-1729
6	4983	99.0	4664	34	US-09-783-514A-1729
7	4983	99.0	4668	21	US-09-396-087-4309
8	4983	99.0	4668	21	US-09-396-970-7559
9	4983	99.0	4668	21	US-09-397-424-5019
10	4983	99.0	4668	21	US-09-397-424A-5019
11	4983	99.0	4668	22	US-09-432-241A-3760
12	4983	99.0	4668	22	US-09-434-737-1278
13	4983	99.0	4668	36	US-09-850-118-1278
14	4983	99.0	6831	1	PCT-US02-27777-49
15	4983	99.0	6831	1	PCT-US02-27777A-49
16	4983	99.0	6871	49	US-10-170-235-39761
17	4971.5	98.8	5371	1	PCT-US02-14570-3
18	4971.5	98.8	5371	36	US-09-851-673-3
19	4971.5	98.8	5371	62	US-10-755-889-489
20	4971.5	98.8	5371	64	US-10-887-553A-199
21	4971.5	98.8	5371	114	US-60-440-068-489
22	4971.5	98.8	5371	116	US-60-469-757-489
23	4970	98.7	6314	102	US-60-324-185-23812
24	4953	98.4	6315	90	US-60-213-359-5981
25	4953	98.4	6316	97	US-60-278-232-7384
26	4942	98.2	4527	34	US-09-786-797B-25
27	4942	98.2	4527	65	US-10-921-707-25
28	4942	98.2	4527	82	US-60-131-321-21
29	4916	97.7	6315	86	US-60-172-360-22693
30	4601	91.4	2845	28	US-09-629-469A-18842
31	4601	91.4	2845	65	US-10-917-503-18842
32	3876.5	77.0	4716	1	PCT-US01-01239-843
33	3876.5	77.0	4716	34	US-09-764-902-843
34	3676	73.0	2523	22	US-09-488-725A-613
35	3676	73.0	2523	46	US-10-037-270-290

ALIGNMENTS

RESULT 1

US-09-671-687A-2
 ; Sequence 2, Application US/09671687A
 ; GENERAL INFORMATION:
 ; APPLICANT: WALLACH, David
 ; APPLICANT: KOVALENKO, Andrei
 ; APPLICANT: CANTARELLA, Giuseppe
 ; TITLE OF INVENTION: INHIBITOR OF NF-KB ACTIVATOR
 ; FILE REFERENCE: WALLACH-25
 ; CURRENT APPLICATION NUMBER: US/09/671,687A
 ; CURRENT FILING DATE: 2000-09-28
 ; PRIOR APPLICATION NUMBER: 09/646,403
 ; PRIOR FILING DATE: 2000-09-18
 ; PRIOR APPLICATION NUMBER: IL 126024
 ; PRIOR FILING DATE: 1998-09-01
 ; PRIOR APPLICATION NUMBER: IL 134604
 ; PRIOR FILING DATE: 2000-02-17
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 2
 ; LENGTH: 3715
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (18)..(18)
 ; OTHER INFORMATION: n is either a, c, g, or t.
 ; NAME/KEY: misc feature
 ; LOCATION: (22)..(22)
 ; OTHER INFORMATION: n is either a, c, g, or t.
 ; NAME/KEY: misc feature
 ; LOCATION: (756)..(756)
 ; OTHER INFORMATION: n is either a, c, g, or t.
 ; NAME/KEY: misc feature
 ; LOCATION: (1348)..(1348)
 ; OTHER INFORMATION: n is either a, c, g, or t.
 ; US-09-671-687A-2

Alignment Scores:
 Pred. No.: 0
 Score: 4992.00
 Percent Similarity: 99.58%
 Best Local Similarity: 99.58%
 Query Match: 30
 DB: 4
 Gaps: 4

US-09-671-687A-3 (1-949) x US-09-671-687A-2 (1-3715)

Qy 1 MetSerSerGlyLeuTTPSerGlnGluValThrSerProTyrTTPGluGluArgIle 20
 Db 497 ATGAGTTCAGGCTATGAGCCAGAAAAAGTCTTCCCTACTGGGAAGCGGATT 556
 Qy -21 PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLys 40
 Db 557 TTTTACTTGTCTTCTCAAGATGCGGCTTACAGACAAACAAACAACTCTTAAA 616
 Qy 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60
 Db 617 GTACCGAAGGGAAGTAGGACAGTATTCAAGATCGTCTGTGGGCGATTCAAGGATT 676

Sequence 290, App
 Sequence 290, App
 Sequence 613, App
 Sequence 613, App
 Sequence 8151, App
 Sequence 8151, App
 Sequence 8151, App
 Sequence 2028, App
 Sequence 4185, App
 Sequence 6145, App
 Sequence 6145, App

QY 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla 80
DB 677 CCTTCTGCAAAAGGCAAGAAATCAGATTGGATTAAAAATCTAGACCAACCTCATGCA 736
QY 81 ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu 99
DB 737 GTTCTCTTTGTGATGAANGGATGTTGTAGAGATAAATGAAAGTTTCACAGAGTTACTT 796
QY 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
DB 797 TTGGCAATTTACCAATTTGTGAGAGAGGTTTCAGGCTGTTTAAAAACAGAAACAGACTAAGT 856
QY 120 LysGlyLeuGluIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
DB 857 AAAGGCTCCAAATAGACTGGGCTCTCTGTGAAAGTACAGCTGAGATCTGGGGAGAA 916
QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
DB 917 AAATTTCTGGAGTTGTAGCTTCAGAGACCCTGTTAGCAGAGAGGACAGTCTCCGGA 976
QY 160 IlePhePheGlyValGluLeuLeuGluGlyArgGlyGlnGlyPheThrAspGlyVal 179
DB 977 ATATCTTTGGAGTTGAAATTTCTGGAGAAAGGTCGTGGTCAAGGTTTCACTGACGGGGTG 1036
QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
DB 1037 TACCAAGGAAACAGCTTTTTCAGTGTGATGATGAAGATTGTGGCTGTTGTGCTATGGAC 1096
QY 199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
DB 1097 AAGCTAGAACTCATAGAAGATGATGACCTGCATTCGAAAGTATTACGAGGTCCTGGG 1156
QY 219 AspThrMetGlnValGluLeuProLeuGluIleAsnSerArgValSerLeuLysGly 238
DB 1157 GACACATGCAGCTCGAACTTCTCTTTGGAAATAAATCTCCAGAGTTTCTTTGAAGGGT 1216
QY 239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
DB 1217 CGAGAAACAATAGAAATCTGGAACAGTTATATCTGTGTGTTTTCGCCGAAAGAAAGC 1276
QY 259 LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe 278
DB 1277 TTAGATATTTTGTGTGTGACATGATGATAACCTTATGGCAACTGGGATGGAAGATT 1336
QY 279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
DB 1337 GATGAGTGCANCTTTGTAGTTTGGCTGTGTGTTGAAAGTACAATTTCTATTGCATCAAT 1396
QY 298 AspIleIleProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMetSer 317
DB 1397 GATATCATCCAGAGAGTGTGCGCAGGAAAGGAGGCGCTCCCAAACTTGCTTTATGTCA 1456
QY 318 ArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThrGlySerThr 337
DB 1457 AGAGTGTGTGGGCAAAAGGTTTCATCCAGTCATAATAAACAAGGCTACAGGATCTACC 1516
QY 338 SerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVal 356
DB 1517 TCAGACCTTGGAAATAGAACAGATCTGAATTTATTTATACCTTAAATGGGCTCTCTGTT 1576
QY 357 AspSerGlnProGlnSerLysAsnThrTyrIleAspGluValAlaGluAsp 376
DB 1577 GACTACACCAACCAATCCAAATCAAAAATAACATGTTACATGATGAGTTGCAGAAAGC 1636
QY 377 ProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProProLeuGln 396
DB 1637 CTGCAAAATCTCTTACAGAGATATCTACAGACTTTTGACCGTTCTTTCACCCACTCCAG 1696
QY 397 ProProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeu 416
DB 1697 CCTCTCTGTGAACCTCACTGACCCAGGAGAACAGATTCCACTCTTTTACCATTCAGTCTC 1756

QY 417 ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGln 436
DB 1757 ACCAAGATGCCAATACCAATGGAAGTATTGGCCACAGTCCACTTCTCTGTGACCCAG 1816
QY 437 SerValMetGluGluLeuAsnThrAlaProValGlnGlnSerProProLeuAlaMetPro 456
DB 1817 TCTGTAATGGAAGGCTAAACACTGCACCCGTCACAGAGAGTCCACCCCTTGGCCATGCT 1876
QY 457 ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro 476
DB 1877 CCTGGGAACCTCACATGCTTAGAAGTGGGCTCATTTGGCTGAAGTTAAGAGAGAACCTCTCT 1936
QY 477 PheTyrGlyValIleArgTrrIleGlyGlnProProGlyLeuAsnGlnValLeuAlaGly 496
DB 1937 TTTCTATGGGGTAATCCGTTGGATCGGTGAGCCACAGGACTGAATGAAGTCTCGCTCGA 1996
QY 497 LeuGluLeuLeuAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr 516
DB 1997 CTGGAACCTGGAAGATGAGTGTGCAGGCTGTACGATGGAACCTTCAGAGGACACTCGGTAT 2056
QY 517 PheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSer 536
DB 2057 TTCACCTGTGCCCTGAAGAAGGCGCTGTTGTGAAACTGAAGAGCTGCAGGCTGACTCT 2116
QY 537 ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerIleAlaPhe 556
DB 2117 AGTGTTCATCATTTGCAGCCGTTTCCAATCAGATTGAGCGCTGTAATCTCTTTAGCATTT 2176
QY 557 GlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluGly 576
DB 2177 GGAGGCTACTTAAGTAGTAGTAGAAGAAATACTCCACCAAAATGGAAAAAGAGGC 2236
QY 577 LeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu 596
DB 2237 TTGGAGATAATGATTGGGAAGAGAAAGGCATCCAGGGTCATTACAATCTTTGTACTTA 2296
QY 597 AspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArg 616
DB 2297 GACTCAACCTTATCTGCTTATTTGCTTTAGTTCTGTTCTGGACACTGTGTTACTTAGA 2356
QY 617 ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu 636
DB 2357 CCACAAAGAAAGAACCATGATAGATAATATTAGTGAAACCCCAAGAGCTACTCAGGACAGAA 2416
QY 637 IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg 656
DB 2417 ATTGTTAACTCTGAGAAATATATGATATGTGTGTCACACAAAAATATGAACCTGAGG 2476
QY 657 LysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGlu 676
DB 2477 AAAATATCTGAAAGGTGGAGGCTGCATCAGGATTTACCTCTGAAGAAAAAGATCCTCAG 2536
QY 677 GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArg 696
DB 2537 GAATTTCTGAATATCTGTTTTCATCATATTTAAGGGTAGAACCTTTTGTCTAAAAATAAGA 2596
QY 697 SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu 716
DB 2597 TCAGCAGGTCAAAAGGTACAAAGATTGTTACTTCTATCAAAATTTTATGGAAAAAATAG 2656
QY 717 LysValGlyValProThrIleGlnLeuLeuGluTrpSerPheIleAsnSerAsnLeu 736
DB 2657 AAAGTTGGGCTTCCCAATTCAGAGTGTGTAGATGGTCTTTTATCAACAGATAACCTG 2716
QY 737 LysPheAlaGluAlaProSerCysLeuIleGlnMetProArgPheGlyLysAspPhe 756
DB 2717 AAATTTGCAGAGGACCATCATGCTCATATTATTCAGATGCTCGATTTGGAAAAAGACTTT 2776
QY 757 LysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAsp 776
DB 2777 AAACTATTTAAAAAATTTTCTCTCTGGAATTAATAATAACAGATTTTACTTGAAGAC 2836
QY 777 ThrProArgGlnCysArgIleCysGlyLeuAlaMetTyrGluCysArgGluCysTyr 796

```
Db      2837  ATCCGAGCAGTCCCGGATATGTGGAGGCTTGCATATGATGAGTGTAGAGATCTCTAC 2896
Qy      797  AspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGln 816
Db      2897  GACGATCCGGACATCTCAGCTGGAAATCAAGCAGTCTTTGTATAAACCTGCAACACTCAA 2956
Qy      817  ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAsp 836
Db      2957  GTCCACCTTCATCCGGAAGAGGCTGAATCATATAATATAACCCAGTGTCACTTCCCAAGAC 3016
Qy      837  LeuProAspTirAspTirPargHisGlyCysIleProCysGlnAsnMetGluLeuPheAla 856
Db      3017  TTACCCGACTGGGACTGGGACACGCGCTGCATCCCTTGCAGAAATATGAGATTATTGGCT 3076
Qy      857  ValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSer 876
Db      3077  GTTCTCTGATAGAAACAGCCACTATGTCTTTGTGAGTATGGGAAGACGATTCT 3136
Qy      877  AlaTirLeuPhePheAspSerMetAlaAspArgAspGlyGlnAsnGlyPheAsnIle 896
Db      3137  GCCTGGCTCTCTTTGACAGCATGGCGGATCGGATGGTGGTCAAGATGGCTTCAACATT 3196
Qy      897  ProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLeu 916
Db      3197  CCTCAAGTCACCCCATGCCAGAAAGTAGGAGAGTACTTGAAGATGTCTCTGGAAGACCTG 3256
Qy      917  HisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyr 936
Db      3257  CATTCCTTGACCTCCAGGAGATCCAAAGCTGTGCAGAAAGACTGCTTTGTGATGATAT 3316
Qy      937  MetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
Db      3317  ATGTGATGTATCCAGAGTCCAAACAATGAGTTGTGTACAAA 3355
```

RESULT 2

US-09-522-303-1293
; Sequence 1293, Application US/09522303
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
; FILE OF INVENTION: HUMAN MAMMARY EPITHELIAL LIBRARY
; TITLE REFERENCE: 1600.1086-001
; CURRENT APPLICATION NUMBER: US/09/522, 303
; CURRENT FILING DATE: 2000-03-08
; EARLIER APPLICATION NUMBER: 60/123, 393
; EARLIER FILING DATE: 1999-03-08
; NUMBER OF SEQ ID NOS: 1353
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1293
; LENGTH: 4661
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(4661)
; OTHER INFORMATION: n = A,T,C or G
US-09-522-303-1293

Alignment Scores:

Pred. No.:	0	Length:	4661
Score:	4983.00	Matches:	948
Percent Similarity:	99.48%	Conservative:	0
Best Local Similarity:	99.48%	Mismatches:	1
Query Match:	98.99%	Indels:	4
DB:	23	Gaps:	4

US-09-671-687A-3 (1-949) x US-09-522-303-1293 (1-4661)

```
Qy      1  MetSerSerGlyLeuTirPserGlnGluLysValThrSerProTyrTirPgluGluArgIle 20
Db      231  ATGAGTTTCAGGCTTATGGAGCCAGAAAAAGTCACTTCACCCCTACTGGGAAGAGCGGATT 290
```

```
Qy      21  PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLeuLys 40
Db      291  TTTTACTTCTCTTCTTCAAGAAATGACGCTTACAGACAAAACAAAACACAAAAGCTCTTTAAA 350
Qy      41  ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60
Db      351  GTACCCGAGGAAAGTATAGACAGTATATTTCAAGATCGTTCTGTGGGCAATTCAGGATT 410
Qy      61  ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla 80
Db      411  CCTTCTGCAAAAGGCAAGAAATATCAGATTGGATTAAAAATTTCTAGAGCAACCTCATGCA 470
Qy      81  ValLeuPheValAspGlu---AspValValGluIleAsnGlnLysPheThrGluLeuLeu 99
Db      471  GTTCTCTTGTGTGATGAAAAAGGATGTTGTAGAGATAAATGAAAAATTTACAGAGATTACTT 530
Qy      100  LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
Db      531  TTGGCAATTACCAATTTGTGAGGAGAGGTTTCAGGCTGTTTAAAAACAGAAAACAGACTAAGT 590
Qy      120  LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
Db      591  AAAGGCTTCNAATATAGCTGGGCTGTCTGTGAAGTACAGCTGAGATCTGGGAAGAA 650
Qy      140  LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
Db      651  AAATTTTCTGGAGTTGCTAGCTTCAGAGGACCCCTGTTTAGCAGAGAGACAGTCTCCGGA 710
Qy      160  IlePhePheGlyValGluLeuLeuGluGlyArgGlyGlnGlyPheThrAspGlyVal 179
Db      711  ATATTCTTTGGAGTTGAAATTTGCTGGAGAGAGTCTGGTCAAGGTTTCTACTGACGGGTG 770
Qy      180  TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
Db      771  TACCAAGGAAACAGCTTTTTCAGTGTGATGAAGATTGTGCGTGTGTTGTTGTCATTGGAC 830
Qy      199  LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
Db      831  AAGCTAGAACTCATAGAAGATGATGACACTGCATTTGGAAAGTGATTACGCGAGTCTCTGGG 890
Qy      219  AspThrMetGlnValGlnLeuProLeuGluIleAsnSerArgValSerLeuLysGly 238
Db      891  GACACAAATGCAGGTCGAACTTCTCTTTGGAAATAAATCCAGAGTTCTTTTGAAGGTT 950
Qy      239  GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
Db      951  GGAGAAACAATAGAAATCTGGAACAGTTATATTCTGTGATGTTTTCGCCAGGAAAAGAAAGC 1010
Qy      259  LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTirPaspGlyArgPhe 278
Db      1011  TTAGGATATTTTGTGGTGGACATGGATAACCTATTGGCAACTGGGATGGGAAGATT 1070
Qy      279  AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
Db      1071  GATGGAGTGCAGCTTTGTAGTCTTTCGTGTGTGTTGAAAGTACAAATCTATTGCACATCAAT 1130
Qy      298  AspileleProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMetSer 317
Db      1131  GATATCATCCAGAGAGTGTGACGCAAGGAGGCTCCCAAACTTGCCTTTATGTGCA 1190
Qy      318  ArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThrGlySerThr 337
Db      1191  AGAGGTGTTGGGACAAAGGTTTATCCAGTCATATATAACCAAGGCTACAGGATCTACC 1250
Qy      338  SerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVal 356
Db      1251  TCAGACCTCGAAATAGAAAACAGACTCTGAATATTTTATATACCTTAAATGGGCTCTCTGT 1310
Qy      357  AspSerGlnProGlnSerLysSerLysAsnThrTirTyrIleAspGluValAlaGluAsp 376
Db      1311  GACTCACACCAATCCAAATCAAAAAATACATGGTACATTGATGAAGTTGCAGAGAAC 1370
```

```
QY 377 ProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerProProLeuGln 396
Db 1371 CTGCAAAATCTCTTACAGAGATATCTACAGACTTTTGACCGTCTCTTCCACCACCTCCAG 1430
QY 397 ProProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeu 416
Db 1431 CCTCTCTCTGTGAACCTCACTGACCCACGAGACAGATTCCACTCTTTACCATTCAGTCTC 1490
QY 417 ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGln 436
Db 1491 ACCAAGATGCCCAATACCAATGAAGTATTGGCCACAGTCCACTTCTCTGTGAGCCAG 1550
QY 437 SerValMetGluLeuLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPro 456
Db 1551 TCTGTAATGGAGAGCTAAACACTGCACCCGTCCTCAAGAGAGTCCACCTTGGCCATGCTT 1610
QY 457 ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro 476
Db 1611 CCTGGGAATCACTGGTCTAGAGTGGGCTCATTTGGCTGMAAGTAAAGAGAACCTCTCT 1670
QY 477 PheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGly 496
Db 1671 TTCTATGGGTAAATCCGTTGGATCGGTCAAGCAGGACTGAATGAAGTGTCTCGCTGGA 1730
QY 497 LeuGluLeuLeuAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr 516
Db 1731 CTGGAACTCGAAGATGAGTGTGAGCTGTACGGATGGAACCTTCAGAGGCACTCGGTAT 1790
QY 517 PheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSer 536
Db 1791 TTCACCTGTGCCCTGAAGAAGCGCTGTTTGTGAACCTGAAGAGCTGCAGGCTGACTCT 1850
QY 537 ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe 556
Db 1851 AGSTTTGCATCATTCAGCGGTTTCCATCAGATTGAGCGCTGAACCTCTTTAGCATTT 1910
QY 557 GlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluGly 576
Db 1911 GGAGGCTACTTAAGTGAAGTAGTAGAAGAAATACTCCACCAAAATGGAAGAAAGGC 1970
QY 577 LeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu 596
Db 1971 TTGGAAGATAATGATTGGGAAGAAAGAGCATCCAGGGTCAATACAAATTTCTTTACTTA 2030
QY 597 AspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArg 616
Db 2031 GACTCAACTTAATCTGCTTAATTTGCTTTAGTTCTGTTCTGGACACTGTGTACTAGA 2090
QY 617 ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu 636
Db 2091 CCCAAAGAAAGAACGATGTAGAAATATATAGTGAACCCCAAGAGCTACTGAGGACAGAA 2150
QY 637 IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg 656
Db 2151 ATTGTTAATCTCTCAGAAATATATGATATGTGTGCCACAAAATTTATGAACCTGAGG 2210
QY 657 LysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGlu 676
Db 2211 AAAATCTTGAAGGTGGGCTGATCAGGATTTTACCTCTGAAGAAAGATCTCTGAG 2270
QY 677 GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLysIleArg 696
Db 2271 GAATCTTGAATATCTGTTTCATCATATTTTAAAGGTAGAACCTTTGCTAAAAATAAGA 2330
QY 697 SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu 716
Db 2331 TCAGCAGGTCAAAAGGTACAGATTGTACTTCTATCAAAATTTTATGGAAGAAATGAG 2390
QY 717 LysValGlyValProThrIleGlnLeuLeuGluTyrSerPheIleAsnSerAsnLeu 736
Db 2391 AAAAGTGGCGTTCCCAAAATTCAGCAGTGTGTAGAAATGCTTTTATCAACAGTAACCTG 2450
QY 737 LysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPhe 756
```

```
Db 2451 AAATTTGCAGAGGCACCATCATGTCTGATTATTTCAGATGCTCGATTTCGAAAAGACTTT 2510
QY 757 LysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAsp 776
Db 2511 AAACATATTTAAAAAATTTTCTCTCTGGAATTAATAATAACAGATTACTTTGAAGAC 2570
QY 777 ThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyr 796
Db 2571 ACTCCAGACAGTCCGCGATATGTGGAGGCTTTCGAATGTATGAGTGTAGAGAAATGCTAC 2630
QY 797 AspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGln 816
Db 2631 GAGATCCGAGACTCATCAGCTGGAATAATCAAGCAGTTTGTAAACCTGCAACACTCRA 2690
QY 817 ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAsp 836
Db 2691 GTCCACCTTCATCCGAAGAGGCTGAATCATAAATAATAACCCAGTGTCACTTCCCAAGAC 2750
QY 837 LeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAla 856
Db 2751 TTACCCGACTGGGACTGGAGACACGGCTGCATCCCTTGCAGAAATATGGAGTTATTTGCT 2810
QY 857 ValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspSer 876
Db 2811 GTTCTCTGCATAGAAACAAGCCACTATGTTGCTTTTGTGAAGTATGGGAAGGAGATTCT 2870
QY 877 AlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIle 896
Db 2871 GCCTGCTCTTCTTTCACAGCATGGCCGATCGGATCGGTGTCAGATGGCTTCAACATT 2930
QY 897 ProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLeu 916
Db 2931 CCTCAAGTCAACCCATGCCAGAAAGTAGGAGACTTCTTGAAGATGTCTCTGGAAGACCTG 2990
QY 917 HisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyr 936
Db 2991 CATTCCTTGACTCCAGGAGAAATCCAAGGCTGTGCACGAAGACTGCTTTGTGATGCATAT 3050
QY 937 MetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
Db 3051 ATGTGCATGTACCAGAGTCCACATGAGTTGTACAAA 3089

RESULT 3
US-09-315-788-1729
; Sequence 1729, Application US/09315788 - Abandoned.
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
; TITLE OF INVENTION: HUMAN ASTROCYTE LIBRARY
; FILE REFERENCE: MLN98-19PM
; CURRENT APPLICATION NUMBER: US/09/315,788
; CURRENT FILING DATE: 1999-05-21
; EARLIER APPLICATION NUMBER: 60/086,455
; EARLIER FILING DATE: 1998-05-22
; EARLIER APPLICATION NUMBER: 60/132,067
; EARLIER FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 2346
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1729
; LENGTH: 4664
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(4664)
; OTHER INFORMATION: n = A,T,C or G
US-09-315-788-1729

Alignment Scores: 0 Length: 4664
Pred. No.: 4983.00 Matches: 948
Score:
```

Percent Similarity:	99.48%	Conservative:	0
Best Local Similarity:	99.48%	Mismatches:	1
Query Match:	98.99%	Indels:	4
DB:	20	Gaps:	4
US-09-671-687A-3 (1-949) x US-09-315-788-1729 (1-4664)			
QY	1	MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluGluArgIle	20
Db	230	ATGAGTTTCAGGCTTATGGAGCCCAAGAAAAGTCACTTCACCCCTACTCGGGAAGCGGAT	289
QY	21	PheTyrLeuLeuLeuGlnCysSerValThrAspLysGlnThrGlnLysLeuLeuLys	40
Db	290	TTTTACTTCTCTTCAAGAAGTCAGCGTTACAGACAAAACAAACAAAGCTCCTTAAA	349
QY	41	ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle	60
Db	350	GTACCGAAGGGAATATAGACAGTATATTCAGATCGTCTCTGTGGGCAATTCAGGAT	409
QY	61	ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla	80
Db	410	CCTTCTGCAAAAGGCAAGAAAATCAGATTGGATTAAATAATTCAGAGCAACCTCATGCA	469
QY	81	ValLeuPheValAspGlu---AspValValGluIleAsnGlnLysPheThrGluLeuLeu	99
Db	470	GTCTCTCTTTGTGATGAAAAGGATGTTGTAGAGATAAATGAAAGTTTCAGAGATTACTT	529
QY	100	LeuAlaIleThrAsnCysGluGluAArgPheSerLeuPheLysAsnArgAsnArgLeuSer	119
Db	530	TTGGCAATTTACCAATTTGTGGAGAGGTTCCAGCTGTTTAAATAACAGAAACAGACTAAGT	589
QY	120	LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu	139
Db	590	AAAGGCTCTCAATATAGACGTGGCTGTCTCTGTGAAAGTACAGCTGAGATCTGGGGAAGA	649
QY	140	LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly	159
Db	650	AAATTTCTCGAGTTGTACGCTTTCAGAGGACCCCTGTTAGCAGAGAGGACAGCTCTCCGGA	709
QY	160	IlePhePheGlyValGluLeuLeuGluGluGlyArgGlyGlnGlyPheThrAspGlyVal	179
Db	710	ATATTTCTTGGAGTTGAAATGCTGGGAAGAGGTCGTGGTCAAGGTTTCACTGACGGGGTG	769
QY	180	TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp	198
Db	770	TACCAAGGNAACAGCTTTTTCAGTGTGATGAGATTGTGGCTGTTTGTTCATTGGAC	829
QY	199	LysLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu	218
Db	830	AACTAGAACTCATAGAAGATGATGACACTGTCATTGGAAAGTGATTAACGAGGTCCTGGG	889
QY	219	AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly	238
Db	890	GACAAATGCGAGTCAACTTCTCTTTGGAAATAAATCCAGAGTTTCTTTTGAAGGTT	949
QY	239	GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer	258
Db	950	GGAGAAACAATAGAACTGGAACAGTTATATCTGTGATGTTTGGCAGGAAAAGAAAGC	1009
QY	259	LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe	278
Db	1010	TTAGGATATTTTGTGTGTGACATGATGATAACCTTATGGCAACTGGGATGGAAGATTT	1069
QY	279	AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn	297
Db	1070	GATGGAGTGCAGCTTTGTAGTTTGGCGTGTGTTGAAAGTAGTACAATTTCTATTGGCATCAAT	1129
QY	298	AspIleIleProGluSerValThrGlnGluArgProProLysLeuAlaPheMetSer	317
Db	1130	GATATCATCCAGAGAGTGTGACGAGAAAGAGGCTCTCCAAACTTGTGCTTTATGTCA	1189
QY	318	ArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThrGlySerThr	337

Db	1190	AGAGGTGTTGGGACAAAGTTTCATCCAGTCATAATAAACCAAGGCTACAGGATCTACC	1249
QY	338	SerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVal	356
Db	1250	TCAGACCCCTGGAAATAGAAACAGATCTGAATTTATTTATACCTTAATATGGCTTCTGTT	1309
QY	357	AspSerGlnProGlnSerLysSerLysAsnThrTyrIleAspGluValAlaGluAsp	376
Db	1310	GACTCACACCAACCAATCCAAATCAAAATAACATGATGATGATGATGATGATGATGATG	1369
QY	377	ProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProLeuGln	396
Db	1370	CCTGCAAAATCTCTTACAGAGATATCTACAGATTTTGACCCGTTCTTCCACCACTCCAG	1429
QY	397	ProProProValAsnSerLeuThrGluAsnArgPheHisSerLeuProPheSerLeu	416
Db	1430	CCTCTCTCTGTGAACCTCACTGACCCAGAGAACAGATTCACCTTTTACCATTTCAGTCTC	1489
QY	417	ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGln	436
Db	1490	ACCAAGATGCCAATACCAATGGAAGTATTTGGCCACAGTCCACTTTCTGTGACGCCAG	1549
QY	437	SerValMetGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu	456
Db	1550	TCTGTAAATGGAAGAGCTAAACACTGCACCCGTCAGAGAGAGTCCACCTTGGCCCATGCT	1609
QY	457	ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro	476
Db	1610	CCTGGGAACTCAGATGCTTAGAAGTGGGCTCATTTGGCTGAAGTTAGAGAGAACCTCTCT	1669
QY	477	PheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGly	496
Db	1670	TTCTATGGGGTAATCCGTTGGATCGGTGAGCCAGCAGGACTGAATGAAGTCTCGCTGGA	1729
QY	497	LeuGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu	516
Db	1730	CTGGAATCTGGAAGATGAGTGTGAGGCTGTACGAGTGAACCTTTCAGAGGACCTCGGTAT	1789
QY	517	PheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSer	536
Db	1790	TTTACCTGTGCTTGAAGAGGCGCTGTTTGTGAACTGAAGAGCTCAGGCTGACTCT	1849
QY	537	ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe	556
Db	1850	AGGTTTGCATCAATTCAGCGGTTTCCAAATCAGATTGAGCGCTGTAACCTTTCAGCATTT	1909
QY	557	GlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluGly	576
Db	1910	GGAGGCTACTTAAAGTAGTAGTAGAGAAAATACTCCACCAAAATGGAAGAAAGGAGGC	1969
QY	577	LeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu	596
Db	1970	TTGGAGATTAATGTTGGAAAGAAAGAGGATCCAGGCTCATTAACAATTTCTTGTACTTA	2029
QY	597	AspSerThrLeuPheCysLeuPheAlaPheSerValLeuAspThrValLeuLeuArg	616
Db	2030	GACTCAACCTTATCTGCTTATTTGCTTTAGTTTCTGTTCTGAGACACTGTGTTACTTGA	2089
QY	617	ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu	636
Db	2090	CCCAAGAAAAGAACGATGTAGAATATATATAGTGAACCCCAAGAGCTACTGAGGACAGAA	2149
QY	637	IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg	656
Db	2150	ATTGTTAATCCTCTGAGATATATGAGATATGTTGTGTCACCAAAATTTATGAACTGAGG	2209
QY	657	LysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGlu	676
Db	2210	AAATACTTGAAGGTTGGAGGCTGCATCAGGATTTTACCTCTGAAGAAAAGATCTCTGAG	2269
QY	677	GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArg	696
Db	2270	GAATTTCTTGAATATTTCTGTTTTCATCATATTTTAAAGGCTAGAACCTTTGCTAAAAATA	2329

```
QY 697 SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu 716
Db 2330 TCAGCAGGTCAAAGGTACAAAGATTGTTACTTCTATCAAAATTTTATGGAAAAATGAG 2389
QY 717 LysValGlyValProThrIleGlnLeuLeuGlnTyrSerPheIleAsnSerAsnLeu 736
Db 2390 AAAGTTGGCGTCCCAACATTCAGCAGTGTAGAGTGTCTTTTATCAACAGTAACCTG 2449
QY 737 LysPheAlaGluAlaProSerCysLeuIleGlnMetProArgPheGlyLysAspPhe 756
Db 2450 AAATTTGCAGAGCACCATCATCTCTGATTTATTCAGATGCTCGATTTGAAAAAGACTTT 2509
QY 757 LysLeuPheLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAsp 776
Db 2510 AAACATATTTAAAAAATTTTCTCTCTGGAATTAATAACAGATTTACTTTGAAGAC 2569
QY 777 ThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyr 796
Db 2570 ACTCCAGACAGTCCCGGATATGTGGAGGCTTGCAATGTATGAGTGTAGAGAAATGCTAC 2629
QY 797 AspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGln 816
Db 2630 GACGATCCGACATCTCAGCTGGAAAAATCAAGCAGTTTGTAAAAACCTGCAACACTCAA 2689
QY 817 ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAsp 836
Db 2690 GTCCACCTTCATCCGAAGAGGCTGAAATCATATAATAAATCAACAGTGTCACTTCCCAAAGAC 2749
QY 837 LeuProAspTyrAspTyrPheHisGlyCysIleProCysGlnAsnMetGluLeuPheAla 856
Db 2750 TTACCCGACTGGGACTGGAGACACGCTGCTATCCCTTGGCAGAAATATGGAGTTATTTGCT 2809
QY 857 ValLeuCysIleGlnThrSerHisTyrValAlaPheValLysTyrGlyLysAspSer 876
Db 2810 GTTCTCTGATAGAAACAGACCACTATGTGTTTGTGAGTATGGAGAGGAGCATTTCT 2869
QY 877 AlaTyrLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIle 896
Db 2870 GCCTGGCTCTCTTTGACAGCATGGCGGATCGGATGGTGGTTCAGAAATGGCTTCAACATT 2929
QY 897 ProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLeu 916
Db 2930 CCTCAAGTCACCCCATGCCAGAGTAGGAGAGTACTTGAAGATGCTCTGGAAGACCTG 2989
QY 917 HisSerLeuAspSerArgIleGlnGlyCysAlaArgLeuLeuCysAspAlaTyr 936
Db 2990 CATTCCTGGACTCCAGGAGATCCCAAGGCTGTGCACAGAGACTGCTTTGTGATGCATAT 3049
QY 937 MetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
Db 3050 ATGTGCATGTACAGAGTCCACATGAGTTGTACAAA 3088
```

RESULT 4

```
US-09-315-788A-1729
; Sequence 1729, Application US/09315788A - Abdn
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
; TITLE OF INVENTION: HUMAN ASTROCYTE LIBRARY
; FILE REFERENCE: 1600, 1019-002
; CURRENT APPLICATION NUMBER: US/09/315, 788A
; CURRENT FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 09/315, 788
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 60/086, 455
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/132, 067
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 2346
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1729
```

```
; LENGTH: 4664
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)....(4664)
; OTHER INFORMATION: n = A, T, C or G
US-09-315-788A-1729
```

Alignment Scores:

```
Pred. No.: 0 Length: 4664
Score: 4983.00 Matches: 948
Percent Similarity: 99.48% Conservative: 0
Best Local Similarity: 99.48% Mismatches: 1
Query Match: 98.99% Indels: 4
DB: 20 Gaps: 4
```

US-09-671-687A-3 (1-949) x US-09-315-788A-1729 (1-4664)

```
QY 1 MetSerSerGlyLeuTyrPheSerGlnGluLysValThrSerProTyrTyrGluGluArgIle 20
Db 230 ATGAGTTTCAGGCTTATGGAGCCCAAGAAAAAGTCACCTTCACTTCTGGAAGAGCGGATT 289
QY 21 PheTyrLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLys 40
Db 290 TTTTACTTGTCTTCTCAAGAAATGCAGGTTACAGACAAACAAACACAAAGCTCTTAAA 349
QY 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60
Db 350 GTACCCAGGGAAGTATAGACAGTATATTCAGATCGTCTGTGGGCGATTCAGGATT 409
QY 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla 80
Db 410 CCTTCTGAAAAGCGCAAGAAAAATCAGATTGGATTAATAATTTAGAGCAACCTCATGCA 469
QY 81 ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu 99
Db 470 GTTCTCTTGTGTTGATGAAAAGGATGTTGTAGAGATAATGAAAAGTTTACAGAGATTACTT 529
QY 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
Db 530 TTGGCAATTTACCAATTTGTGAGGAGAGGTTGAGCTGTTTAAAAACAGAAACAGACTAAGT 589
QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
Db 590 AAAGGCTCCAAATAGACGTGGGCTGTCTGTGAAAGTACAGCTGAGATCTGGGGAAGAA 649
QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
Db 650 AAATTTCTCGAGTTGTACGCTTCAGAGGACCCCTGTGTAGCAGAGAGACAGTCTCCGGA 709
QY 160 IlePhePheGlyValGluLeuLeuGluGlyArgGlyGlnGlyPheThrAspGlyVal 179
Db 710 ATATTTCTTGGAGTTGAATTTGCTGGAAGAGGCTGTGCTCAAGGTTTCACTGACGGGGTG 769
QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
Db 770 TACCAAGGGAACAGCTTTTTCAGTGTGATGAAGATTGTGGCGTGTGTTGTCATTGGAC 829
QY 199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
Db 830 AAGCTAGAACTCATAGAAGATGATGACACTGCACTTGGAAAGTATTACCGAGGCTCTGGG 889
QY 219 AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly 238
Db 890 GACACAAATGCAGGTCGAACCTTCTCTTTGGAATAAACTCCAGAGTTTCTTTGAAGTT 949
QY 239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
Db 950 GGAGAAACATAGAAATCTGGAACAGTTATATTCGTGATGTTTTCGCCAGAAAGAAAGC 1009
QY 259 LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTyrAspGlyArgPhe 278
```


Db 1010 TTAGGATATTTTGGTGGGACATGGATACCCCTATTGGCAACTGGGATGGAAGATTT 1069
QY 279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
Db 1070 GATGGAGTGCAGCTTTGTAGTTTGGCTGTGTTGAAAGTACAATTTCTATTGGACATCAAT 1129
QY 298 AspIleIleProGluSerValThrGlnGluArgProProLysLeuAlaPheMetSer 317
Db 1130 GATATCATCCAGAGAGTGTGACGCGAGGAAGAGGAGCCCTCCCAAACTTGCCTTTATGTCA 1189
QY 318 ArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThrGlySerThr 337
Db 1190 AGAGGTGTGGGACCAAGGTTTCATCCAGTCATATAATAACCAAGGCTACAGGATCTACC 1249
QY 338 SerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVal 356
Db 1250 TCAGACCCCTGGAAATAGAAACAGATCTGAATATTTTATACCTTAAATGGGTCTTCTGTT 1309
QY 357 AspSerGlnProGlnSerLysSerLysAsnThrTyrIleAspGluValAlaGluAsp 376
Db 1310 GACTCAACAACCAATCCAAATCAAAATAATACATGTTGATGAAAGTTGCAGAAGAC 1369
QY 377 ProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerProProLeuGln 396
Db 1370 CCTGCAAAATCTCTTACAGAGATATCTACAGACTTTTGACCGTTCTTCACCACCACCTCCAG 1429
QY 397 ProProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeu 416
Db 1430 CCTCTCTCTGTGNACTCCTGACACCGAGACAGATTCACCTCTTTACATTCAGTCTC 1489
QY 417 ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGln 436
Db 1490 ACCAAGATGCCAATACCAATGAAAGTATTGGCCACAGTCCACTTCTCTGTGAGCCAG 1549
QY 437 SerValMetGluLeuLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPro 456
Db 1550 TCTGTAATGGAGAGCTAAACACTGCACCCGCTCAAGAGAGTCCACCCCTTGGCCATGCCT 1609
QY 457 ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro 476
Db 1610 CCTGGGAACCTCATGGCTAGAAAGTGGGCTCAATGGCTGAAGTGAAGGAACCCCTCT 1669
QY 477 PheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGly 496
Db 1670 TTCTATGGGGTAAATCCGTTGGATCGGTGAGCCAGGACTGAATGAAGTCTCGTGGTA 1729
QY 497 LeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr 516
Db 1730 CTGGAACTGGAAAGATGAGTGTGAGGCTGTACGGATGGAACCTTCAGAGGCACTCGGTAT 1789
QY 517 PheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSer 536
Db 1790 TTCACCTGTGCTGAAAGAGGCGCTGTTGTGAAACTGAAGAGCTGCAGGCGCTGACTCT 1849
QY 537 ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe 556
Db 1850 AGGTTTGCATCATGTGACGCGGTTTCCAATCAGATTGAGGCGCTGAACCTTTAGCAATTT 1909
QY 557 GlyGlyTyrLeuSerGluValValGluAsnThrProProLysMetGluLysGluGly 576
Db 1910 GGAGGCTACTTAAGTGAAGTAGTAGAAGAAATACTCCACCAAAATGGAAGAAAGAGGC 1969
QY 577 LeuGluIleMetIleGlyLysLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu 596
Db 1970 TTGGAGATATGATTGGGAAGAGAAAGGCAATCAGGGTCAATTACAAATCTCTTGTACTTA 2029
QY 597 AspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArg 616
Db 2030 GACTCAACCTTATCTGCTTATTGCTTTTGTCTTTAGTTCTGTTCTGGACACTGTGTACTTGA 2089
QY 617 ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu 636
Db 2090 CCCAAGAAAGAAACGATGTAGATATTTATAGTGAAACCCCAAGAGCTACTGAGGACAGAA 2149

QY 637 IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg 656
Db 2150 ATTGTTAAATCCTCTGAGAATATATGATGATGTGTGCCACAAAAATTTATGAAACTGAGG 2209
QY 657 LysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGlu 676
Db 2210 AAAATACTTTGAAAAGGTGAGGCTGCATCAGGATTTTACCTCTCAAGAAAAAAGATCCCTGAG 2269
QY 677 GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArg 696
Db 2270 GAATTTCTTGAATATCTGTTTCATCATATTTTAAAGGGTAGAACCTTTGCTTAAAAATAAGA 2329
QY 697 SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu 716
Db 2330 TCAGCAGGTCAAAAGGTACAGATTTGTTACTTCTATCAAAATTTTATGGAAAAAATGAG 2389
QY 717 LysValGlyValProThrIleGlnLeuLeuGluTrpSerPheIleAsnSerAsnLeu 736
Db 2390 AAAGTTGGCGTTTCCCACAATTCAGCAGTTGTTTAAAGATGGTCTTTTATCAACAGTAACCTG 2449
QY 737 LysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPhe 756
Db 2450 AAATTTGCAGAGCCACCATCATGTCATGATTTTTCAGATGCTCGATTTGGAAAAAGACTTT 2509
QY 757 LysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAsp 776
Db 2510 AAACATTTTAAAAAAATTTTCTCTCTCGAATTAATAATACAGATTTTACTTGAAGAC 2569
QY 777 ThrProArgGlnCysArgIleCysGlyLeuAlaMetTyrGluCysArgGluCysTyr 796
Db 2570 ACTCCACAGACAGTCCGGATATGTGAGGCGCTTGCAATGTATGAGTGTAGAGAATGCTAC 2629
QY 797 AspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGln 816
Db 2630 GACGATCCCGACACTCAGCTGAAAAAATCAACGAGTTTGTAAAAACCTGCAACACTCAA 2689
QY 817 ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAsp 836
Db 2690 GTCCACTTCATCCGAAGAGGCTGAATCATATATATACCCAGTGTCACTTCCCAAGAC 2749
QY 837 LeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAla 856
Db 2750 TTACCCGACTGGGACTGGAGACACGCTGCATCCCTTCCAGAAATATGGAGTTATTGCT 2809
QY 857 ValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSer 876
Db 2810 GTTCTCTGCATAGAAAACAAGCCACTATGTTGCTTTTGTGAAGTATGGGAAGGACGATCT 2869
QY 877 AlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIle 896
Db 2870 GCTGCTCTCTTTTTCAGCATGGCGGATGGGATGGTGGTTCAGAAATGGCTTCAACATT 2929
QY 897 ProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLeu 916
Db 2930 CCTCAAGTCAACCCATGCCAGAGTAGGAGAGTACTTGAAGATGTCTCTGGAAGACCTG 2989
QY 917 HisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyr 936
Db 2990 CATTCCTTGGACTCCAGGAATCCAAAGGCTGTGCACGAAGACTGCTTTGTGATGCATAT 3049
QY 937 MetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
Db 3050 ATGTGATGTACCAGAGTCCAAACATGAGTTGTGACAAA 3088

RESULT 5

US-09-783-514-1729
; Sequence 1729, Application US/09783514
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
; HUMAN ASTROCYTE LIBRARY

- Adm


```
QY 577 LeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu 596
Db 1970 TTGGAGATAATGATGGGAAGAAAGGCATCCAGGGTCATTACAAATCTTGTTACTTA 2029

QY 597 AspSerThrLeuPheCysLeuPheAlaPheSerValLeuAspThrValLeuLeuArg 616
Db 2030 GACTCAACCTTATCTGCTTATTTGCTTTAGTCTGTTCTGGACACTGTGTTACTTAGA 2089

QY 617 ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu 636
Db 2090 CCCAAAGAAAGAACCATGATAGTAATATTATAGTGAACCCCAAGCTACTGAGACAGAA 2149

QY 637 IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg 656
Db 2150 ATTGTTAATCTCTGAGATATATGATATGTTGTGCCACAAAAATTTATGAACCTGAGG 2209

QY 657 LysIleLeuGluLysValGluAlaLaSerGlyPheThrSerGluGluLysAspProGlu 676
Db 2210 AAAATACTTGAAGAGTGGAGGCTGCTCAGGATTTACCTCTGAAGAAAGATCCTGAG 2269

QY 677 GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArg 696
Db 2270 GAATCTTGNATATTTCTGTTCTATCATATTTTAAAGGGTAGAACCTTTGCTAAAAATAAGA 2329

QY 697 SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu 716
Db 2330 TCAGCAGGTCAAAAGGTACAAGATTGTTACTTCTATCAATTTTATGGAAGAAAAATGAG 2389

QY 717 LysValGlyValProThrIleGlnLeuLeuGluTyrSerPheIleAsnSerAsnLeu 736
Db 2390 AAAGTTGGGCTTCCCAATTCACAGAGTTGTTAGAATGGTCTTTTATCAACAGTAACCTG 2449

QY 737 LysPheAlaGluAlaProSerCysLeuIleGlnMetProArgPheGlyLysAspPhe 756
Db 2450 AAATTTGCAGAGCACATCATGCTCTGATTTATTTACAGATGCTCGATTTGGAAGAACTTT 2509

QY 757 LysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAsp 776
Db 2510 AAACATTTAAAAAATTTTCTCTCTGGAATTAATAATACAGATTTACTTTGAAGAC 2569

QY 777 ThrProArgGlnCysArgIleCysGlyLeuAlaMetTyrGluCysArgGluCysTyr 796
Db 2570 ACTCCACAGACAGTCCGGATATGTGGAGGGCTTGCATATGATGATGATGAGTAAGTCTAC 2629

QY 797 AspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGln 816
Db 2630 GACGATCCGGACATCTCAGCTGGAAATAATCAAGCAGTTTGTAAAAACCTCGCAACACTCAA 2689

QY 817 ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAsp 836
Db 2690 GTCCACCTTCATCCGAAGAGCTGATCATATATATATACCATGCTGCTCTCCCAAGAC 2749

QY 837 LeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAla 856
Db 2750 TTACCCGACTGGGACTGGACACAGCGCTGCATCCCTTCCAGAAATATGAGATTTATTTGCT 2809

QY 857 ValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSer 876
Db 2810 GTTCTCTGCATAGAAACCAACCATCATGTTGCTTTTGTGAAGTATGGGAAGGACGATCT 2869

QY 877 AlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlnAsnGlyPheAsnIle 896
Db 2870 GCCTGGCTCTTCTTTGACATGCGCCGATCGGATGGTGGTTCAGAAATGGCTTCAACATT 2929

QY 897 ProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLeu 916
Db 2930 CCTCAAGTCACCCCATGCCAGAGTAGGAGAGTACTTGAAGATGCTCTCTGGAAGACCTG 2989

QY 917 HisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgIleLeuCysAspAlaTyr 936
Db 2990 CATTCCTTGGACTCCAGAGATCCCAAGGCTGTGCAGGAAGACTGCTTTGTGATGCATAT 3049
```

```
QY 937 MetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
Db 3050 ATGTGCATGTACCAGAGTCCAAATGAGTTTGTACAAA 3088

RESULT 6
US-09-783-514A-1729 - Adm
; Sequence 1729, Application US/09783514A
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
; TITLE OF INVENTION: HUMAN ASTROCYTE LIBRARY
; FILE REFERENCE: 1600.1019-002
; CURRENT APPLICATION NUMBER: US/09/783,514A
; CURRENT FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 09/315,788
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 60/086,455
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/132,067
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 2346
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1729
; LENGTH: 4664
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(4664)
; OTHER INFORMATION: n = A,T,C or G
US-09-783-514A-1729
```

```
Alignment Scores:
Pred. No.: 0 Length: 4664
Score: 4983.00 Matches: 948
Percent Similarity: 99.48% Conservative: 0
Best Local Similarity: 99.48% Mismatches: 1
Query Match: 98.99% Indels: 4
DB: 34 Gaps: 4
```

US-09-671-687A-3 (1-949) x US-09-783-514A-1729 (1-4664)

```
QY 1 MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluGluArgIle 20
Db 230 ATGAGTTTCAGCTTATGGAGCCCAAGAAAAGTCACTTCACTTCTGGAAGACGGGATT 289

QY 21 PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLys 40
Db 290 TTTTACTTGTCTTCTCAAGAAATGCAGCGTTACAGACAAACAAACAAAGCTCCTTAAA 349

QY 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60
Db 350 GTACCCGAAGGGAAGTAGTAGGACAGTATATTTCAAGATCGTTCTGTGGGCGATTCAAGGATT 409

QY 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla 80
Db 410 CCTTCTGCAAGCAAGCAAGAAAATCAGATTGGATTAAAAATTTCTAGACCACTCATGCA 469

QY 81 ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu 99
Db 470 GTTCTCTTCTTCTGATGAAAGGATGTTGTAGAGATAAATGAAAAGTTTCAACAGATTACTT 529

QY 100 LeuAlaIleThrAsnCysGluLysArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
Db 530 TTGGCAATTACCAATTTGTGTAGGAGAGGTTTCAAGCTGTGTTAAAAACAGAAAACAGACTAAGT 589

QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
Db 590 AAAGGCTTCCAAATAGACGTGGGCTGCTCTGTGAAAGTACAGCTGAGATCTGGGGAAGAA 649

QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
```

Db 650 AAATTTCTCTGGAGTTGTACGCTTCAGAGGACCCCTGTGTAGCAGAGGACAGTCTCCGGA 709
Qy 160 IlePhePheGlyValGluLeuLeuGluGluGlyArgGlyGlnGlyPheThrAspGlyVal 179
Db 710 ATATTCTTTGGAGTTGAATTCGTGGAGAGGCTGTGGTCAAGGTTTCTACTGACGGGGTG 769
Qy 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
Db 770 TACCAGGGAAACAGCTTTTTCAGTGTGATGAAGATTGTGGCTGTTTGTGTGATTGGAC 829
Qy 199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
Db 830 AAGCTAGAACTCATAGAAGATGATGACACTGCAATTGGAAAGTATTACCGAGTCTCTGG 889
Qy 219 AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly 238
Db 890 GACACAATGCAGTTCGAATCTCTCTTTGGAAATAAATCCAGAGTTTCTTTGAAGTT 949
Qy 239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
Db 950 GGAGAAACAATAAGATCTGGAACAGTTATTCTGTGATGTTTGTCCAGGAAAGAAAGC 1009
Qy 259 LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe 278
Db 1010 TTAGGATATTTTGTGTGTGACATGATTAACCTATTGGCAACTGGGATGGAAGATT 1069
Qy 279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
Db 1070 GATGGAGTGCAGTTTGTAGTTTGTGCTGTGTTGAAAGTACAATTCTATTGGACATCAAT 1129
Qy 298 AsnIleIleProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMetSer 317
Db 1130 GATATCATCTCCAGAGAGTGTGACGAGGAAAGAGGCCTCCCAAACTTGCCTTTATGTCA 1189
Qy 318 ArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThrGlySerThr 337
Db 1190 AGAGTGTGTGGGACAAAGGTTTCATCCAGTCATATAAACCAGGCTACAGATCTACC 1249
Qy 338 SerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVal 356
Db 1250 TCAGACCTCTGGAAATAGAAACAGATCTGAATATTTTATACCTTAAATGGTCTCTGTT 1309
Qy 357 AspSerGlnProGlnSerLysSerLysAsnThrTyrTyrIleAspGluValAlaGluAsp 376
Db 1310 GACTCACACCAACCAATCCAAATCAAAAATACATGTTACATGATCAAGTTGCAGAAGAC 1369
Qy 377 ProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProProLeuGln 396
Db 1370 CCTGCAAAATCTTTACAGAGATATCTACAGACTTTTGACCGTTCCTTCAACACACCTCCAG 1429
Qy 397 ProProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeu 416
Db 1430 CTCTCTCTGTGAACTCACTGACCCAGGAGACAGATTCCACTCTTTACATTCACTCTC 1489
Qy 417 ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLysSerLeuSerAlaGln 436
Db 1490 ACCAAGATGCCCAATACCAATGGAAGTATTGGCCACAGTCCACTTCTCTGTGAGCCAG 1549
Qy 437 SerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPro 456
Db 1550 TCTGTAATCGAAGAGCTAAACACTGACCCCGTCCAAAGAGAGTCCACCTTGGCCATGSCCT 1609
Qy 457 ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro 476
Db 1610 CTGGGAATCTCAATGGTCTAGAAAGTGGGCTCATTTGGCTGAAGTTAAGGAGAACCTCTCT 1669
Qy 477 PheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGly 496
Db 1670 TTCTATGGGTAAATCCGTTGGATCGGTGACGCCACAGGACTGAATGAAGTGTCTCGTGA 1729
Qy 497 LeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr 516
Db 1730 CTGGAACCTGGAAGATGAGTGTGAGGCTGTACGGATGGAACCTTTCAGAGGCACTCGGTAT 1789

Qy 517 PheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSer 536
Db 1790 TTCACCTGTGCTCGAAGAAAGGCGCTGTTTGTGAAACTGAGAGCTCGAGCCCTGACTCT 1849
Qy 537 ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe 556
Db 1850 AGTGTTCATCATGTGACGCGGTTTCCAACTCAGATTGAGCGCTGTAACTCTTTAGCATTT 1909
Qy 557 GlyGlyTyrLeuSerGluValValGluAsnThrProProLysMetGluLysGluGly 576
Db 1910 GGAGGCTACTTAAGTGAAGTAGTAGAAGAAATACTCCACCAAAATGGAAAAAGAGGC 1969
Qy 577 LeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu 596
Db 1970 TTGGAGATAATGATTTGGGAAGAAAGGATCCAGGCTCATTAACAATTTCTTTGTACTTA 2029
Qy 597 AspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArg 616
Db 2030 GACTCAACCTTATTCGCTTATTTGCTTTTAGTCTCTGTTCTGGACACTGTGTTACTTGA 2089
Qy 617 ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu 636
Db 2090 CCCAAAGAAAGAACCGATGTAGAATATTTAGTGAAACCCCAAGAGCTACTGAGCAGAA 2149
Qy 637 IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg 656
Db 2150 ATTGTTAAATCTCTGAGATATATGATATGTGTGTGCCACAAAAATTTATGAAACTCGAG 2209
Qy 657 LysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGlu 676
Db 2210 AAAATCTTTGAAAAGTGGAGGCTGCATCAGGATTTTACCTCTGAAGAAAAAGATCCTGAG 2269
Qy 677 GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArg 696
Db 2270 GAATTTCTTGATATTTCTGTTTCATCATATTTTAAAGGTAGAACCTTTGCTTAAAAATAAGA 2329
Qy 697 SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu 716
Db 2330 TCAGCAGGTCAAAAGGTACAAATTTGTTACTTCTCATCAAAATTTTATGGAAAAAATGAG 2389
Qy 717 LysValGlyValProThrIleGlnGlnLeuLeuGluTrpSerPheIleAsnSerAsnLeu 736
Db 2390 AAAAGTTGGGCTTCCCAACTTCAGCAGTGTGTAGAATGGTCTTTTATCAACAGTAACCTG 2449
Qy 737 LysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPhe 756
Db 2450 AAATTTGCAGAGGACCATCATGTCTGATTTATTCAGATGCTCGATTTTGGAAAAAGACTTT 2509
Qy 757 LysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAsp 776
Db 2510 AAATTAATTTAAAAAATTTTCTCTCTCGAATTAATAATAACAGATTTTACTTTGAAGAC 2569
Qy 777 ThrProArgGlnCysArgIleCysGlyLeuAlaMetTyrGluCysArgGluCysTyr 796
Db 2570 ACTCCAGACAGTGCCTGATATGTGGAGGCTTGCATGTATGTAGTGTAGAGATGTCTAC 2629
Qy 797 AspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGln 816
Db 2630 GACGATCCGGACATCTCAGCTGGAAAAATCAAGCAGTTTGTAAAAACCTGCAACACTCAA 2689
Qy 817 ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAsp 836
Db 2690 GTCCACTTCATCCGAGAGGCTGAATCATATAATAACCCAGTGTCTACTTCCCAAGAC 2749
Qy 837 LeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAla 856
Db 2750 TTACCCGACTGGGACTGGAGACAGGCTGCATCCCTTGCAGAAATATGGAGTTATTGCT 2809
Qy 857 ValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSer 876
Db 2810 GTTCTCTGTAGAAAAACGCACTATGTTGCTTTTGTGAAGTATGGGAAGGACGATTTCT 2869

QY 877 AlaTrrLeuPheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIle 896
 DB 2870 GCCTGGCTCTTTTGACAGCATGGCCGATCGGATGGGTCAGAAATGGCTTCAACATT 2929
 QY 897 ProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLeu 916
 DB 2930 CCTCAAGTCACCCCATGCCAGAGTAGAGAGTACTTGAAGATGTCCTTGGAGACCTCG 2989
 QY 917 HisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyr 936
 DB 2990 CAFTCTTGAGACTCCAGGAGATCCAAAGCTGTGCACGAAGACTGCTTTGTGATGCATAT 3049
 QY 937 MetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
 DB 3050 ATGTGCATGTACCAGAGTCCAAATAGTGTGTACAAA 3088

RESULT 7
 US-09-396-087-4309 - Abdn
 ; Sequence 4309, Application US/09396087
 ; GENERAL INFORMATION:
 ; APPLICANT: Gearing, David P.
 ; APPLICANT: Holtzman, Douglas A.
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
 ; TITLE OF INVENTION: PULMONARY ARTERY SMOOTH MUSCLE LIBRARY
 ; FILE REFERENCE: MLN98-39pM
 ; CURRENT APPLICATION NUMBER: US/09/396,087
 ; CURRENT FILING DATE: 1999-09-14
 ; EARLIER APPLICATION NUMBER: 60/100,260
 ; EARLIER FILING DATE: 1998-09-14
 ; EARLIER APPLICATION NUMBER: 60/107,226
 ; EARLIER FILING DATE: 1998-11-05
 ; EARLIER APPLICATION NUMBER: 60/131,810
 ; EARLIER FILING DATE: 1999-04-30
 ; NUMBER OF SEQ ID NOS: 5220
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 4309
 ; LENGTH: 4668
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(4668)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-396-087-4309

Alignment Scores:
 Pred. No.: 0 Length: 4668
 Score: 4983.00 Matches: 948
 Percent Similarity: 99.48% Conservative: 0
 Best Local Similarity: 99.48% Mismatches: 1
 Query Match: 98.99% Indels: 4
 DB: 21 Gaps: 4

US-09-671-687A-3 (1-949) x US-09-396-087-4309 (1-4668)

QY 1 MetSerSerGlyLeuTrrSerGlnGluLysValThrSerProTyrTrpGluArgIle 20
 DB 230 ATGAGTTCAGGCTTATGGAGCCGAAAGAAAGTCACTTCCCTTCTGGGAGACGGATT 289
 QY 21 PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLys 40
 DB 290 TTTTACTTGCTTCAAGAATGCAGGTTTACAGACAAACAAACAAAGCTCCTTAA 349
 QY 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisArgIle 60
 DB 350 GTACCAAGGGAAGTAGACAGATATATTCAAGATCGTTCTGTGGGGCAATCAAGATT 409
 QY 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGlnProHisAla 80
 DB 410 CCTTCTGCAAAAGGCAAGAAATACAGATTGGATTAAATTTAGAGCAACCTCATGCA 469
 QY 81 ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeu 99

DB 470 GTTCTCTTTGTTGATGATAAAGGATGTTGTAGAGATAAATGAAAGTTCACAGAGTTACTT 529
 QY 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
 DB 530 TTGGCAATTTACCAATTTGTGAGAGAGGTTTCAGCTGTTTAAAAACAGAAACAGACTAAGT 589
 QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
 DB 590 AAAGGCTTCCAAATAGACGTGGCTGCTCTGTGAAAGTACAGCTGAGATCTGGGGAAGAA 649
 QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
 DB 650 AAATTTCTCTGGAGTTGTACGCTTCAGAGACCCCTGTAGCAGAGAGGACAGTCTCCGGA 709
 QY 160 IlePhePheGlyValGluLeuGluGlyArgGlyGlnGlyPheThrAspGlyVal 179
 DB 710 ATATTCTTTGGAGTTGAATTTGCTGGAGAGGTCGTGGTCAAGGTTTCACTGACGGGGTG 769
 QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
 DB 770 TACCAAGGGAACAGCTTTTTCAGTGTGATGAAGATTGTGGCTGTTGTTCATTGGAC 829
 QY 199 LysLeuLeuLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
 DB 830 AAGCTAGAACTCATAGAATGATGACACTGCAATGGAAAGTGAATACGCGAGGTCCTGGG 889
 QY 219 AspThrMetGlnValGluLeuProLeuGluIleAsnSerArgValSerLeuLysGly 238
 DB 890 GACACANTGCGAGTTCGAATCTCTCTTTTGGAAATAAATCCAGAGTTTCTTTGAAAGGTT 949
 QY 239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
 DB 950 GGAGAAACAATAAGATCTGGAACAGTTATATCTGTGATGTTTTCAGGAGAAAGAAC 1009
 QY 259 LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe 278
 DB 1010 TTAGGATATTTTGTGGTGGACATGGATAACCTATTGGCACTGGGATGGAGATT 1069
 QY 279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
 DB 1070 GATGGAGTGCAGCTTTGTAGTTTGTGTTGTTGAAAGTACAATTTCTATTGCACTCAAT 1129
 QY 298 AspIleIleProGluSerValThrGlnGluArgProProLysLeuAlaPheMetSer 317
 DB 1130 GATATCATCCAGAGAGTGTGACGAGGAAAGGAGGCTCCCAAACTTGCTTTATGTCA 1189
 QY 318 ArgGlyValGlyAspLysGlySerSerSerHisAsnLysProLysAlaThrGlySerThr 337
 DB 1190 AGAGGTGTTGGGGACAAAGGTTTCATCCAGTCAATAAATAAACCAAGGCTACAGGATCTACC 1249
 QY 338 SerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVal 356
 DB 1250 TCAGACCTTGGAAATAGAAACAGATCTGAATATTTTATACCTTAAATGGGTCTCTGTT 1309
 QY 357 AspSerGlnProGlnSerLysSerLysAsnThrTrrTyrIleAspGluValAlaGluAsp 376
 DB 1310 GACTCACACCAACATCCAAATCAAAAAATACATGTTGATGATGAAGTTGCAGAGAAC 1369
 QY 377 ProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProLeuGln 396
 DB 1370 CCTGCAAAATCTCTTACAGAGATATCTACAGACTTTTGACCGTTTCCACCACCACTCCAG 1429
 QY 397 ProProValAsnSerLeuThrGluAsnArgPheHisSerLeuProPheSerLeu 416
 DB 1430 CCTCTCTCTGTGAACCTACCTGACCCAGAACAGATTCCACTCTTTACCAATTGAGTCTC 1489
 QY 417 ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGln 436
 DB 1490 ACCAAGATGCCAATACCAATGGAGTATTGGCCACAGTCCACTTCTCTGTGAGCCAG 1549
 QY 437 SerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPro 456
 DB 1550 TCTGTAATGGAAGAGCTAAACACTGCACCCGTCAGAGAGTCCACCTTGGCCATGCCT 1609

```
QY 457 ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro 476
Db 1610 CTGGGAACCTCACATGCTCTAGAAAGTGGGCTCATTTGGCTGAAGTAAAGGAGAACCCCTCCT 1669
QY 477 PheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGly 496
Db 1670 TTCATAGGGGTAACTCGTTGGATCGGTACGCCACAGGACTGAATGAAGTGGCTCGCTGGA 1729
QY 497 LeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr 516
Db 1730 CTGGAACCTGGAAGATCAGTGTGCGAGCTGTACGGATGGAACCTTCAGAGGCACCTCGGTAT 1789
QY 517 PheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSer 536
Db 1790 TTCACCTGTGCCCTGAAGAAGGCGCTTTGTGTGAACCTGAAGAGCTGCAGGCCCTGACTCT 1849
QY 537 ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe 556
Db 1850 AGTTTGGCATCATTTGCGAGCGGTTTCCAATCAGATTGAGCGCTGTAACCTTTTAGCATTT 1909
QY 557 GlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluGly 576
Db 1910 GGAGGCTACTTAAGTGAAGTAGTAGAAGAAATACTCCACCAAAATGGAAAAAGAGGC 1969
QY 577 LeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu 596
Db 1970 TTGGAGATTAATGATTGGGAAGAGAAAGGCATCCAGGGTCATTACAAATCTCTGTACTTA 2029
QY 597 AspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArg 616
Db 2030 GACTCAACCTTAATCTGCTTATTTGCTTTTAGTTCTGTTCTGGACACTGTGTACTTAGA 2089
QY 617 ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu 636
Db 2090 CCCAAAGAAAGAACGATGTAGAAATATATATGTGAAACCCAGAGCTACTGAGACAGAA 2149
QY 637 IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg 656
Db 2150 ATTGTTAATCCTGAGATATATGATGATATGTGTGCCACAAAAATTTATGAACCTGAGG 2209
QY 657 LysIleLeuGluLysValGluAlaIleAspGlyPheThrSerGluGluLysAspProGlu 676
Db 2210 AAAATACTTTGAAAGGTGGAGGCTGCATCAGGATTTTACCTCTCGAAGAAAGATCTCTGAG 2269
QY 677 GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArg 696
Db 2270 GAATCTTGAATATTTCTGTTTCATCATATTTTAAAGGGTAGAACCTTTTGCTAAAAAATAAGA 2329
QY 697 SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu 716
Db 2330 TCAGCAGGTCAAAAGGTACAGATTTGTACTTCTATCAATTTTATGGAAAAAATGAG 2389
QY 717 LysValGlyValProThrIleGlnGlnLeuLeuGluTrpSerPheIleAsnSerAsnLeu 736
Db 2390 AAAAGTTGGCGTTCCCAACATTCAGCAGTTGTGTAGAAATGGTCTTTTATCAACAGTAACCTG 2449
QY 737 LysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPhe 756
Db 2450 AAAATTTGCAGAGGCACCATCATGCTCTGATTATTTAGATGCTCGATTTTGGAAAAAGACTTT 2509
QY 757 LysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAsp 776
Db 2510 AAACATATTTAAAAAATTTTCTCTCTGGAATTAATATAACAGATTTTACTTTGAAGAC 2569
QY 777 ThrProArgGlnCysArgIleCysGlyLeuAlaMetTyrGluCysArgGluCysTyr 796
Db 2570 ACTCCAGACAGTCCGGATATGTGGAGGCTTGCAATGATGATGATGATGATGATGATGATG 2629
QY 797 AspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGln 816
Db 2630 GACGATCCGACATCTCAGCTGGAAAAATCAAGCAGTTTTTGTAAAAACCTGCAACACTCAA 2689
```

```
QY 817 ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAsp 836
Db 2690 GTCCACCTTCATCCGAAGAGGCTGAATATAAAATATAACCCAGTGTCTACTTCCCAAGAC 2749
QY 837 LeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAla 856
Db 2750 TTACCCGACTGGGACTGGAGACAGGCTGCATCCCTTGGCCAGAAATATGGAGTTATTGCT 2809
QY 857 ValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSer 876
Db 2810 GTTCTCTGCATAGAAAACAGCCCATATGTTGCTTTTGTGAAGTATGGGAAGACGATTTCT 2869
QY 877 AlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlnAsnGlyPheAsnIle 896
Db 2870 GCTGGCTCTTCTTTTGACATGCGCGATGCGGATGGTGCAGATGGCTTCAACATT 2929
QY 897 ProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLeu 916
Db 2930 CCTCAAGTCAACCCATGCCAGAAAGTAGGAGAGTACTTTGAAGATGTCTCTGGAGACCTG 2989
QY 917 HisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyr 936
Db 2990 CATTCCTTGGACTCCAGGAGAAATCCAAGGCTGTGCACGAGACTGCTTTGTGTGCATAT 3049
QY 937 MetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
Db 3050 ATGTGCATGTACCAGAGTCCAAACAATGAGTTTGTACAAA 3088

RESULT 8
US-09-396-970-7559
; Sequence 7559, Application US/09396970 -> Node
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
; FILE REFERENCE: MLN98-40Pa
; CURRENT APPLICATION NUMBER: US/09/396,970
; CURRENT FILING DATE: 1999-09-14
; EARLIER APPLICATION NUMBER: 60/100,293
; EARLIER FILING DATE: 1999-09-14
; NUMBER OF SEQ ID NOS: 8756
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7559
; LENGTH: 4668
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(4668)
; OTHER INFORMATION: n = A,T,C or G
US-09-396-970-7559

Alignment Scores:
Pred. No.: 0 Length: 4668
Score: 4983.00 Matches: 948
Percent Similarity: 99.48% Conservative: 0
Best Local Similarity: 99.48% Mismatches: 1
Query Match: 98.99% Indels: 4
DB: 21 Gaps: 4

US-09-671-687A-3 (1-949) x US-09-396-970-7559 (1-4668)
QY 1 MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluArgIle 20
Db 230 ATGAGTTTCAGGCTTATGGAGCCAAAGAAAAAGTCACTTACCCCTACTGGGAAGAGCGGATT 289
QY 21 PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLys 40
Db 290 TTTTACTTGTCTTCTCAAGAAATGCGGCTTACAGACAAACAAACAAAGCTCTTAAA 349
QY 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60
```

Db 350 GTACCGAAGGAGTATAGACAGTATATTCAAGATCGTTCTGTGGGCAATTCAGGATT 409
Qy 61 ProSerAlaIysGlyLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla 80
Db 410 CCTTCTGCAAAAGGCAAGAAATACAGATTGGATTAAAAATTCAGAGCAACCTCATGCA 469
Qy 81 ValLeuPheValAspGlu---AspValValGluIleAsnGlnLysPheThrGluLeuLeu 99
Db 470 GTTCTCTTTGTGATGAAAGGATGTTGTAGAGATAAATGAAAGTTTCAGAGTACTT 529
Qy 100 LeuAlaIleThrAsnCysGluGluAtrgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
Db 530 TTGGCAATTTACCAATTTGTGAGAGAGTTTCAGCTGTTTAAACAGAAACAGACTAAGT 589
Qy 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
Db 590 AAAGGCTCCAAATAGACGTGGCTGCTCTGTGAAAGTACAGCTGAGATCTGGGGAAGAA 649
Qy 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
Db 650 AAATTTCTTGGAGTTGTAGCTTTTCAGAGACCCCTGTTTACAGAGAGACAGTCTCCGGA 709
Qy 160 IlePhePheGlyValGluLeuLeuGluGluGlyAtrgGlyGlnGlyPheThrAspGlyVal 179
Db 710 ATATTTCTTTGGAGTTGAATTTCTGGAGAGAGTGTGTCAGAGTTTCACTGACGGGGTG 769
Qy 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
Db 770 TACCAAGGGAACAGCTTTTTCAGTGTGATGAAGATTGTGGCTGTTTGTTCATTTGGAC 829
Qy 199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
Db 830 AAGCTAGAACTCATAGAGATGATGACACTGCATTGGAAAGTGATTTACGACGAGTCTCTGG 889
Qy 219 AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly 238
Db 890 GACACAATGCAGTCCGAATCTCTCTCTTGGAAATAAATCCAGAGTTTCTTTTGAAGGTT 949
Qy 239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
Db 950 GGGAAACAATAGAACTCGAACAGTTATATCTGTGATGTTTTCAGGAAAGAAAGAC 1009
Qy 259 LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe 278
Db 1010 TTAGGATATTTTGTGTGTGGACATGATGATAACCTATTGGCAACTGGGATGGAAGATT 1069
Qy 279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
Db 1070 GATGGAGTGCAGCTTTGTAGTTTGTGGTGTGTTGAAAGTACAAATCTTATTTGCAATCAAT 1129
Qy 298 AspIleIleProGluSerValThrGlnGluAtrgArgProProLysLeuAlaPheMetSer 317
Db 1130 GATATCATCCAGAGAGTGTGACGCAAGAAAGAGGCCCTCCCAAACTTGTCTTTATGTCA 1189
Qy 318 ArgGlyValGlyAspLysGlySerSerSerHisAsnLysProLysAlaThrGlySerThr 337
Db 1190 AGAGGTGTTGGGACAAAGTTTCATCCAGTCATAATAAACCAAGGCTACAGGATCTACC 1249
Qy 338 SerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVal 356
Db 1250 TCAGACCCCTGGAAATAGAAACAGATCTGAATATATTTATACCTTAAATGGGTCTTCTGTT 1309
Qy 357 AspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGluValAlaGluAsp 376
Db 1310 GACTCAACACCAATCCAAATCAAAAAATACATGTTGATGATGATGATGATGATGATGATGAT 1369
Qy 377 ProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProProLeuGln 396
Db 1370 CCTGCAAAATCTTACAGAGATATCTACAGACTTTGACCGTCTTCCACCACTCCAG 1429
Qy 397 ProProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeu 416
Db 1430 CCTCTCTCTGTGAACCTCACTGACCACCGAGAACAGATTCCACTCTTTTACCATTCCTC 1489

Qy 417 ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGln 436
Db 1490 ACCAAGATGCCCAATACCAATGGAAAGTATTGGCCACAGTCCACTTTCTCTGTGAGCCAG 1549
Qy 437 SerValMetGluGluLeuLeuAsnThrAlaProValGlnGlnSerProProLeuAlaMetPro 456
Db 1550 TCTGTAAATGGAAAGAGCTAAACACTGCACCCGCTCCAAGAGAGTCCACCTTTGGCCATGCT 1609
Qy 457 ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro 476
Db 1610 CTTGGGAACTCACATGGTCTAGAAAGTGGGCTCAATTTGGCTGAAGTTAAGAGAAACCTCTCT 1669
Qy 477 PheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGly 496
Db 1670 TTCTATGGGGTAAATCCGTTGGATCGGTGAGCCACAGGACTGAATGAAGTGTCTGCTGGA 1729
Qy 497 LeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr 516
Db 1730 CTGGAACTGGAAAGATGAGTGTGAGGCTGTACGGATGGAACTTTCAGAGGCACTCGGTAT 1789
Qy 517 PheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSer 536
Db 1790 TTCACCTGTGCTTGAAGAGGCGCTGTTTGTGAAACTGAAGAGCTGCAGGCTGACTCT 1849
Qy 537 ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe 556
Db 1850 AGTTTTCATCATTTGCAGCCGTTTCCAAATCAGATTGAGCGCTGTAACTCTTTAGCATTT 1909
Qy 557 GlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluGly 576
Db 1910 GGAGGCTACTTAAGTGAAGTAGTAGAAGAAATACTCCACCAAAATGGAAAAAGAGGC 1969
Qy 577 LeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu 596
Db 1970 TTGGAGATAATGATTCGGAGAGAAAGAGCATCCAGGGTCAATACAAATCTTGTACTTA 2029
Qy 597 AspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArg 616
Db 2030 GACTCAACCTTATCTGCTTATTTGCTTTTGTGTTCTGTTCTGGACACTGTGTACTTAGA 2089
Qy 617 ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu 636
Db 2090 CCAAGAGAAAGAACGATGTAGAATATATAGTGAACCAACCAAGAGCTACTGAGACAGAA 2149
Qy 637 IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg 656
Db 2150 ATGTAAATCTCTGAGAAATATATGATATGTGTGTGCCACAAAAATTTATGAAACTGAGG 2209
Qy 657 LysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGlu 676
Db 2210 AAATTAATTTGAAAGGTGGAGGCTGCATCAGGATTTACCTCTGAAGAAAAAGATCTCTGAG 2269
Qy 677 GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArg 696
Db 2270 GAATTTCTTGAATATTTCTGTTTCATCATATTTTAAAGGTAGACCTTTGTCTAAAAATAAGA 2329
Qy 697 SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu 716
Db 2330 TCAGCAGGTCAAAAGGTACAAGATTGTACTTCTATCAAAATTTTATGGAATAAATGAG 2389
Qy 717 LysValGlyValProThrIleGlnLeuLeuGluTrpSerPheIleAsnSerAsnLeu 736
Db 2390 AAAGTTGGGCTTCCCAATTCAGCAGTGTGTGAATGGTCTTTTATCAACAGTAACTGT 2449
Qy 737 LysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPhe 756
Db 2450 AAATTTGCAGAGGCAACCATCATGCTGATTTATTTCAGATGCTCTGATTTGGAAGACTTT 2509
Qy 757 LysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAsp 776
Db 2510 AAATATTTAAAAAATTTTTTCTCTCTGGAATTAATATATAACAGATTTTACTTGAAGAC 2569

QY 777 ThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyr 796
DB 2570 ACTCCAGACAGTCCCGGATATGTGAGGGCTTGCAATGATGAGTGTAGAGTAATCTAC 2829
QY 797 AspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGln 816
DB 2630 GACGATCCGGACATCTCAGCTGGAAATATCAACGAGTTTGTGTAACCTGCAACACTCAA 2689
QY 817 ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAsp 836
DB 2690 GTCCACCTTTCATCCGAAGAGGCTGAATCATATAATATAAACCAGCTGTCACTTCCCAAGAC 2749
QY 837 LeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAla 856
DB 2750 TTACCCGACTGGGACTGGGACACGCTGCTATCCCTTGCAGAAATATGGAGTATTGTGCT 2809
QY 857 ValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSer 876
DB 2810 GTTCTCTGATAGAAACAGCCACTATGTGCTTTTGTGAAGTATGGGAAGGAGGATTCT 2869
QY 877 AlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIle 896
DB 2870 GCTGCTCTTCTTTCACAGCATGGCGATCGGATGGTGGTCAAGTATGGGATGGCTTCAACATT 2929
QY 897 ProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLeu 916
DB 2930 CCTCAAGTCACCCATGCCAGAGTAGGAGAGTACTTGAAGATGTCTCTGGAAGACCTG 2989
QY 917 HisSerLeuAspSerArgIleGlnGlyCysAlaArgLeuLeuCysAspAlaTyr 936
DB 2990 CATTCCTTGAGCTCCAGGAGATCCAAAGCTGTGCAGAGACTGCTTTGTGATGATAT 3049
QY 937 MetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
DB 3050 ATGTGATGTACAGAGTCCAAACATGAGTTGTACAAA 3088

RESULT 9

US-09-397-424-5019
; Sequence 5019, Application US/09397424 — *Abdu*
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: Pan, Yang
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
; FILE REFERENCE: MLN98-45pm
; CURRENT APPLICATION NUMBER: US/09/397,424
; CURRENT FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: 60/100,469
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/106,454
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 60/107,252
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: 60/132,100
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 5379
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5019
; LENGTH: 4668
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(4668)
; OTHER INFORMATION: n = A,T,C or G
US-09-397-424-5019

Alignment Scores:

Pred. No.:	0	Length:	4668
Score:	4983.00	Matches:	948
Percent Similarity:	99.48%	Conservative:	0
Best Local Similarity:	99.48%	Mismatches:	1
Query Match:	98.99%	Indels:	4

DB: 21 Gaps: 4
US-09-671-687A-3 (1-949) x US-09-397-424-5019 (1-4668)
QY 1 MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluGluArgIle 20
DB 230 ATGAGTTTCAGGCTTATGGAGCCCAAGAAAAGTCACTTCACTTCTGGGAAGAGCGGATT 289
QY 21 PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLeuLys 40
DB 290 TTTTACTTGTCTTCTCAAGAAATGCAGCGTTACAGACAAACAAACACAAAAGCTCTCTAAA 349
QY 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60
DB 350 GTACCGAAGGAGTATAGGACAGTATATTCAAGATCGTCTGTGGGCGCATTCAGGATT 409
QY 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla 80
DB 410 CCTTCTGCAAAAGCAAGAAAATCAGATTGGATTAAAATTCTTAGAGCAACCTCATGCA 469
QY 81 ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu 99
DB 470 GTTCTCTTTCATGAAAAGGATGTTGTAGAGATAAATGAAAAGTTTCACAGAGTTACTT 529
QY 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
DB 530 TTGGCAATTACCAATTGTGAGGAGAGGTTTCAGGCTGTTTAAAAACAGAAAACAGACTAAGT 589
QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
DB 590 AAAGGCTTCCAAATAGACGTGGGCTGCTCTGTGAAAGTACAGCTGAGATCTGGGGAAGAA 649
QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
DB 650 AAATTTCTGGAGTTGTACGCTTCAGAGGACCCCTGTAGCAGAGAGACAGCTCTCCGA 709
QY 160 IlePhePheGlyValGluLeuLeuGluGluGlyArgGlyGlnGlyPheThrAspGlyVal 179
DB 710 ATATTCTTTGGAGTTGAATTTGCTGGAAGAGGTCGTGCTCAAGGTTTTCACCTGACGGGGTG 769
QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
DB 770 TACCAAGGGAAACAGCTTTTTCAGTGTGATGAAGATTGTGCGGTGTTTGTTCATTTGAC 829
QY 199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
DB 830 AAGCTAGACTCATAGAGATGATGACACTGCATTGGAAGTGTATTACGAGGTCCTGGG 889
QY 219 AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly 238
DB 890 GACACAATGCAGGTCGAACTTCTCTTGGAAATAAATCCAGAGTTTCTTTTGAAGGTT 949
QY 239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
DB 950 GGAGAAACAAATAGAAATCTGGAACAGTTATATTCTGTATGTTTTCGCCAGGAAAGAAAGC 1009
QY 259 LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTyrAspGlyArgPhe 278
DB 1010 TTAGGATATTTTGTGTGTGGACATGGATAACCTATTGGCAACTGGGATGGAGATT 1069
QY 279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
DB 1070 GATGGAGTCAGCTTTGTAGTTTTCGTGTGTGTGAAAGTACAATTCTATTTCACATCAAT 1129
QY 298 AspIleIleProGluSerValThrGlnGluArgProProLysLeuAlaPheMetSer 317
DB 1130 GATATCATCCAGAGAGTGTGACGAGAAAGGAGGCTCTCCAAACTTGCCTTTATGTCA 1189
QY 318 ArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThrGlySerThr 337
DB 1190 AGAGGTGTGGGACAAAGGTTTCATCCAGTCATATAAACCAGGCTACAGGATCTAC 1249
QY 338 SerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVal 356

Db	1250	TCAGACCTCGAAATAGAAA	CAGATCTGAAATATATTTTATACCTTAAATGGGTCTCTGTT	1309
Qy	357	AspSerGlnProGlnSerLysSerLysAsnThrTyrTyrIleAspGluValAlaGluAsp	376	
Db	1310	GACTCACAACCAACATCCAAATCAAAAATACATGGTACATTGATGAAGTTGCAGAGAC	1369	
Qy	377	ProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProProLeuGln	396	
Db	1370	CCTGCAAAATCTCTTACAGAGATATCTACAGACTTTGACCGTCTTCCACCACTCCAG	1429	
Qy	397	ProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeu	416	
Db	1430	CCTCCTCTGTGAACTCACTACACCGAGAACAGATTCCACTCTTTACCATTCAGTCTC	1489	
Qy	417	ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGln	436	
Db	1490	ACCAAGATGCCCAATACCAATGGAAAGTATTGGCCACAGTCCACTTCTCTGTGAGCCAG	1549	
Qy	437	SerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPro	456	
Db	1550	TCTGTANTGGNAGACTTAACACTGCACCGTCCAAGAGAGTCCACCTTGGCCATGCCT	1609	
Qy	457	ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro	476	
Db	1610	CCTGGGAACCTCACATGGTCTAGAAAGTGGGCTCAATGGCTGAAGTTAAGGAGAACCCCTCCT	1669	
Qy	477	PheTyrGlyValIleArgTyrIleGlyGluProProGlyLeuAsnGluValLeuAlaGly	496	
Db	1670	TTCTATGGGGTAATCCGTGGATCGGTGAGCCACGACACTGAATGAAGTGTCTGCTGGA	1729	
Qy	497	LeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr	516	
Db	1730	CTGGAACTGGNAGATGAGTGTGACGGCTGTACGATGGNACCTTCAGAGGCATCGGTAT	1789	
Qy	517	PheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSer	536	
Db	1790	TTCACTGTGCCCTGAAGAGGCGCTGTTGTGAAACTGAAGAGCTCGAGCGCTGACTCT	1849	
Qy	537	ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe	556	
Db	1850	AGGTTTGCATCATTTGACCGCGTTTCCAATCAGATTGAGCGCTGTAACTCTTTAGCAATT	1909	
Qy	557	GlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluGly	576	
Db	1910	GGAGCTACTTAAGTGAAGTAGTAGAAGAAATACTCCACAAAATGGAAAGAAGGC	1969	
Qy	577	LeuGluLeuMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu	596	
Db	1970	TTGAGATAATGATTGGGAAGAAAGGCATCCAGGCTCATTAACAATCTTGTGTACTTA	2029	
Qy	597	AspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArg	616	
Db	2030	GACTCAACCTTATTTCTGCTTATTTGGCTTTTAGTTCTGTCTGGACACTGTGTACTTAGA	2089	
Qy	617	ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu	636	
Db	2090	CCCAAGAAAAGACGATGTAGAATATTTATAGTGAACCCCAAGAGCTACTGTGAGACAGAA	2149	
Qy	637	IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg	656	
Db	2150	ATTGTTTAATCTCTGAGAATATATGGATATGTGTGTGCCACAAAATTTATGAAACTCAGG	2209	
Qy	657	LysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGlu	676	
Db	2210	AAAAATACCTTGAAAAGGTGGAGGCTGCATCAGGATTTTACCTCTGAGAGAAAAGATCCTGAG	2269	
Qy	677	GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArg	696	
Db	2270	GAATTCCTTGAATATTCGTGTTTCATCATATTTTAAAGGTAGAACCTTTGCTTAAAAATAGA	2329	
Qy	697	SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu	716	

RESULT 10

```

US-09-397-424A-5019
; Sequence 5019, Application US/09397424A - Aden
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: Pan, Yang
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
; TITLE OF INVENTION: HUMAN OSTEOBLAST LIBRARY
; FILE REFERENCE: MLN98-45pm
; CURRENT APPLICATION NUMBER: US/09/397,424A
; CURRENT FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: 60/100,469
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/106,454
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 60/107,252
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: 60/132,100
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 5379
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5019
; LENGTH: 4668

```

```

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(4668)
; OTHER INFORMATION: n = A,T,C or G
; US-09-397-424A-5019

Alignment Scores:
Pred. No.: 0 Length: 4668
Score: 4983.00 Matches: 948
Percent Similarity: 99.48% Conservative: 0
Best Local Similarity: 99.48% Mismatches: 1
Query Match: 98.99% Indels: 4
DB: 21 Gaps: 4

US-09-671-687A-3 (1-949) x US-09-397-424A-5019 (1-4668)

QY 1 MetSerGlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluGluArgIle 20
DB 230 ATGAGTTGAGCTTATGGAGCCCAAGAAAGTCACTTCCCTTCTGGGAGAGCGGATT 289
QY 21 PheTyrLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLys 40
DB 290 TTTTACTTCTCTTCAAGAATGCAGGTTTACAGACAAACAAACAAAGCTCTTAAA 349
QY 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60
DB 350 GTACCAAGGGAAGTAGACAGTATATTCAAGATCGTTCTGTGGGCGATTCAAGGATT 409
QY 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGlnProHisAla 80
DB 410 CCTTCTGCAAAAGGCAAGAAATACAGATTGGATTAAATTTAGACCAACTCATGCA 469
QY 81 ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu 99
DB 470 GTTCTCTTTGATGAAAGGATGTTGTAGATATAATGAAAGTTTCACAGATTACTT 529
QY 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
DB 530 TTGGCAATTTACCAATTTGTGAGGAGAGGTTTCAGCTGTGTTTAAACAGAACAGACTAAGT 589
QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
DB 590 AAAGGCTTCCAAATAGACGTGGCTCTCTGTGAAAGTACAGCTGAGATCTGGGGAAGA 649
QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
DB 650 AATTTCTTGGAGTTGTACGCTTTCAGAGGACCCCTGTTAGCAGAGGACAGCTCCCGGA 709
QY 160 IlePhePheGlyValGluLeuLeuGluGluArgGlyArgGlyGlnGlyPheThrAspGlyVal 179
DB 710 ATATCTTTGGAGTTGATTTGCTGGAGAGAGGTCGTGGTCAAGGTTTCACTGACGGGTG 769
QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
DB 770 TACCAAGGGAACACAGCTTTTTCAGTGTGATGAAGATTGTGGCGTGTGTTGTTGATTGGAC 829
QY 199 LysLeuLeuLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
DB 830 AAGCTAGAACTCTAGAAGATGATGACACTGCAATTGGAAAGTGAATACGCAGGTCTCTGGG 889
QY 219 AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly 238
DB 890 GACCAATGCGAGTCAAGCTTCTCTCTTGGAAATAACTCCAGAGTTCTTTTGAAGGTT 949
QY 239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
DB 950 GGAAGAACCAATAGAACTCGAACAGTATATTCTGTGATGTTTGTCCAGGAAAGAAAGC 1009
QY 259 LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe 278
DB 1010 TTAGGATATTTTGTGGTGGACATGATTAACCCCTATTGGCACTGGGATGGAAGATT 1069

```

```

QY 279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
DB 1070 GATGGAGTGCAGCTTTGTAGTTTTCGTGTGTTGAAAGTACAAATCTATTTCACATCAAT 1129
QY 298 AspIleIleProGluSerValThrGlnGluArgProProLysLeuAlaPheMetSer 317
DB 1130 GATATCATCCAGAGAGTGTGCGCAGAGAAAGAGGCTCCCAAACCTTTCCTTTATGTCA 1189
QY 318 ArgGlyValGlyAspLysGlySerSerSerHisAsnLysProLysAlaThrGlySerThr 337
DB 1190 AGAGGTGTTGGGGACAAAGGTTTCATCCAGTCATATAAACCAAGGCTTACAGGATCTACC 1249
QY 338 SerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVal 356
DB 1250 TCAGACCCTGGAATAGAAACAGATCTGAATTTATTTATACCTTAAATGGGTCTTCTGTT 1309
QY 357 AspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGluValAlaGluAsp 376
DB 1310 GACTCACACCAACCAATCCCAATCAAAAATACATGTTGATGATGAAGTTGCAGAGAC 1369
QY 377 ProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProLeuGln 396
DB 1370 CCTGCAAAATCTCTTACAGAGATATCTACAGACTTTGACCGTTCTTACACCACTCCAG 1429
QY 397 ProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeu 416
DB 1430 CCTCCTCTGTGAACCTCAGCCACCGAGAACAGATTCCACTCTTTTACCATTTCAGTCTC 1489
QY 417 ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGln 436
DB 1490 ACCAAGATGCCCAATCAATGGAAGTATTGGCCACAGTCCACTTCTCTGTAGCCAG 1549
QY 437 SerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPro 456
DB 1550 TCTGTAATGGAAGAGCTAAACACTGCACCCGTCACAGAGAGTCCACCTTGGCCATGCT 1609
QY 457 ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro 476
DB 1610 CCTGGAACTCACAATGCTTAGAAGTGGGCTCATTTGGCTGAAGTTAAGGAAACCTCCT 1669
QY 477 PheTyrGlyValIleArgTyrIleGlyGlnProGlyLeuAsnGluValLeuAlaGly 496
DB 1670 TTCTATGGGTAAATCCGTTGGATCGGTGCGTCCAGCCAGGACTGAATGAAGTCTCGCTGA 1729
QY 497 LeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr 516
DB 1730 CTGGAATCGAAGATGAGTGTGAGGCTGTACGATGGAACCTTCAGAGGCACTCGGTAT 1789
QY 517 PheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSer 536
DB 1790 TTCACCTGTGCCCTGAAGAGGCGCTGTTTGTGAACCTGAAGAGCTGCAGGCCCTGACTCT 1849
QY 537 ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe 556
DB 1850 AGTGTGCACTCATTTGCAGCGGTTTCCAAATCAGATTGAGCGCTGTAATCTTCTTAGCATTT 1909
QY 557 GlyGlyTyrLeuSerGluValValGluLeuAsnThrProProLysMetGluLysGluGly 576
DB 1910 GGAGGCTTACTTAAAGTAGTAGAAGAAATACTCCACCAAAATGGAAAAAGAGAGC 1969
QY 577 LeuGluIleMetIleGlyLysLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu 596
DB 1970 TTGGAGATATGATTGGGAAGAGAAAGGCAATCAGGCTCATTAACATTTCTTTACTTAA 2029
QY 597 AspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArg 616
DB 2030 GACTCAACCTTATCTGCTTATTGCTTTTGTAGTTCTGTTCTGGACACTGTGTACTTAGA 2089
QY 617 ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu 636
DB 2090 CCAAAAGAAAGAACGATAGATAATATTATAGTGAAACCCCAAGAGCTACTAGGAGACAAA 2149

```

```
QY 637 IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg 656
Db 2150 ATGTGTAATCTCTGAGATATATGATATGTGTGCCACAAAATATGAAACTGAG 2209
QY 657 LysIleLeuGluLysValGluAlaSerGlyPheThrSerGluGluLysAspProGlu 676
Db 2210 AAATACTTGAAGGTGGGGCTGATCAGGATTTACCTCTGAGAGAAAAGATCTCTGAG 2269
QY 677 GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArg 696
Db 2270 GAATCTTGAATATTCTGTTTCATCATATTTAAAGGCTAGAACCTTTGCTTAAATAAGA 2329
QY 697 SerAlaGlyClnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu 716
Db 2330 TCAGCAGGTCAAAAGGTACAAGATTGTTACTTATCAAAATTTTATGGAATAAATGAG 2389
QY 717 LysValGlyValProThrIleGlnLeuLeuGluTyrPheThrIleAsnSerAsnLeu 736
Db 2390 AAAGTTGGCGTCCCAATTCAGCAGTTGTTAGAATGGTCTTTTATCAACAGTAACTCG 2449
QY 737 LysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPhe 756
Db 2450 AAATTTGCAGAGCCACCATCATCTCTGATTTATTCAGATGCTCGATTTGGAAGAAGCTTT 2509
QY 757 LysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAsp 776
Db 2510 AAACATTTTAAAAAATTTTCTCTCTGGAATTAATAATAACAGATTTTACTTGAAGAC 2569
QY 777 ThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyr 796
Db 2570 ACTCCAGACAGTGCAGGATATCTGAGGGCTTGCATATGATGAGTGTAGAGATGCTAC 2629
QY 797 AspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGln 816
Db 2630 GAGATCCGGACATCTCAGCTGGAAATATCAACAGTTTTTGTAAACCTGCAACACTCAA 2689
QY 817 ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAsp 836
Db 2690 GTCCACCTTCATCCGAAGAGCTGAATATATATAATTAACCCAGTGTCACTTCCCAAGAC 2749
QY 837 LeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAla 856
Db 2750 TTACCCGACTGGGACTGGAGACAGCGCTGCATCCCTTGCAGAAATATGAGATTTATTGCT 2809
QY 857 ValLeuCystIleGlnThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSer 876
Db 2810 GTTCTCTGCATAGAAACAGCCACTATGTGCTTTTGTGAAGTATGGGAAGGACGATCT 2869
QY 877 AlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIle 896
Db 2870 GCCTGGCTCTTCTTGACAGCATGGCCGATCGGATGGTGTGATGATGCTTCAACATT 2929
QY 897 ProGlnValThrProCysProGluValGlyGlyLeuLysMetSerLeuGluAspLeu 916
Db 2930 CCTCAAGTCAACCCCATGCCAGAGTAGGAGAGTACTTGAAGATGCTCTCTGGAAGACCTG 2989
QY 917 HisSerLeuAspSerArgArgIleGlnGlyCysAlaArgLeuLeuCysAspAlaTyr 936
Db 2990 CATTCCTTGACCTCCAGGAGAAATCCAAAGCTGTGCAGAGATGTCTTGTGATGCATAT 3049
QY 937 MetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
Db 3050 ATGTGATGTACAGAGTCCCAACATGAGTTGTATCAA 3088
```

RESULT 11

```
US-09-432-241A-3760
; Sequence 3760, Application US/09432241A - Abdn
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: Hoitman, Douglas A.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
; TITLE OF INVENTION: HUMAN EPIDERMAL KERATINOCYTE LIBRARY
; FILE REFERENCE: 1600.1004001
```

```
; CURRENT APPLICATION NUMBER: US/09/432,241A
; CURRENT FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/106,445
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 60/107,227
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: 60/127,182
; PRIOR FILING DATE: 1999-03-30
; PRIOR APPLICATION NUMBER: 60/132,099
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 5041
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3760
; LENGTH: 4668
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(4668)
; OTHER INFORMATION: n = A,T,C or G
US-09-432-241A-3760
```

```
Alignment Scores:
Pred. No.: 0 Length: 4668
Score: 4983.00 Matches: 948
Percent Similarity: 99.48% Conservative: 0
Best Local Similarity: 99.48% Mismatches: 1
Query Match: 98.99% Indels: 4
DB: 22 Gaps: 4
```

US-09-671-687A-3 (1-949) X US-09-432-241A-3760 (1-4668)

```
QY 1 MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluGluArgIle 20
Db 230 ATGAGTTTCAGGCTTATGGAGCCCAAGAAAAGTCACTTCCCTACTCGGAGAGCGGATT 289
QY 21 PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLeuLys 40
Db 290 TTTTACTTGTCTTCTCAAGAAATGCGGTTACAGACAAACACAAACAAAGCTCTTAAA 349
QY 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60
Db 350 GTACCGAAGGAAAGTAGTAGCAGATATATTCAAGATCGTTCGTGGGCGCATTTCAAGATT 409
QY 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGlnProHisAla 80
Db 410 CCTTCTGCAAAAGGCAAGAAAATCAGATTGGATTTAAAAATTTAGAGCAACCTCATGCA 469
QY 81 ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu 99
Db 470 GTTCTCTTTGTGATGAAAGGATGTTGTAGAGATAAATGAAAGTTTACAGAGTTACTT 529
QY 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
Db 530 TTGGCAATTAACCAATTGTGAGAGAGGTTTCAGCTGCTTTTAAAAACAGAAACAGACTAAGT 589
QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
Db 590 AAAGGCTTCAAAATAGACGTGGGCTGTCTGTGAAAGTACAGCTGAGATCTGGGGAGAGAA 649
QY 140 LysPheProGlyValValArgPheAtgGlyProLeuLeuAlaGluArgThrValSerGly 159
Db 650 AAATTTCTGGAGTTGTACGCTTCAGAGGAGCCCTGTTAGCAGAGACAGACAGTCTCCGGA 709
QY 160 IlePhePheGlyValGluLeuLeuGluGlyArgGlyGlnGlyPheThrAspGlyVal 179
Db 710 ATATTTCTGGAGTTGAATTTGCTGGAAGAGGTCGTGCTCAAGGTTTCACTGACGGGTG 769
QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
Db 770 TACCAAGGGAACAGCTTTTTCAGTGTGATGAAGATTGTGCGCTGTTGTTGTCATTGGAC 829
QY 199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
```

830 AAGCTAGAACTCTATAGAAAGATGATGACACTGCAATGGAAAGTGAATTACGACAGTCTCTGGG 889
219 AspThrMetGlnValGluLeuProProLeuGluLeuAAsnSerArgValSerLeuLysGly 238
890 GACACAATGCAGGTCGAACTTCTCTCTTTGGAAATAAACTCCAGAGTTCTTTTGAAGTT 949
239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
950 GGAGAAACAATAAGAACTCTGGAACAGATTATCTCTGATGTTTGGCAGGAAAAGAAGC 1009
259 LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe 278
1010 TTAGGATATTTGTTGGTGTGGACATGGATTAACCTATTGGCAACTGGGATGGAGATT 1069
279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
1070 GATGGAGTCAGCTTTGTAGTTTGGGTGTGTTGAAAGTACAATTCTATTGGACATCAAT 1129
298 AsnIleIleProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMetSer 317
1130 GATATCATCCAGAGAGTGTGACGAGGAAAGAGGCGCTCCCAAACTTGCCTTTATGTCA 1189
318 ArgGlyValGlyAspLysGlySerSerSerHisAsnLysProLysAlaThrGlySerThr 337
1190 AGAGGTGTTGGGACAAAGGTTTCATCCAGTCATAATAACCAAGGCTACAGGATCTACC 1249
338 SerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVal 356
1250 TCAGACCTCGGAAATAGAACAGATCTGAATTATTTTATACCTTAATGGGTCTCTGT 1309
357 AspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGluValAlaGluAsp 376
1310 GACTCACAAACCAATCCAAATCAAAATAATACATGCTACATTCATGAAAGTTGCAGAAGAC 1369
377 ProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProProLeuGln 396
1370 CCTGCAAAATCTCTTACAGAGATATCTACAGACTTTTGACCGTTCTTCCACCACCACTCCAG 1429
397 ProProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeu 416
1430 CTTCTCTCTGTGACTACTGACCAACCGAGACAGATTCCACTCTTTACCAATTCAGTCTC 1489
417 ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGln 436
1490 ACCAAGATGCCCAATACCAATGGAAGTATTGGCCACAGTCCACTTCTCTGTCTGAGCCAG 1549
437 SerValMetGluLeuLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPro 456
1550 TCTGTAATGGAAGAGCTAAACACTGCACCGCTCCAAAGAGAGTCCACCCCTTGGCCATGCCT 1609
457 ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro 476
1610 CTTGGAACTCAATGGTCTAGAAGTGGGCTCAATGGCTGAAAGTAAAGAGAACCCCTCCT 1669
477 PheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGly 496
1670 TTCTATGGGGTAAATCCGTTGGATCGTTCAGCCACACAGGACTGATGAAAGTCTCGCTGGA 1729
497 LeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr 516
1730 CTGGAACCTGGAAGATGAGTGTGCAGGCTGTACGGATGGAACCTTTCAGAGGCACCTCGGTAT 1789
517 PheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSer 536
1790 TTCCACTGTGCGCTGAAGAGGCGCTGTTGTGAAACTGAAGAGCTGCAGGCGCTGACTCT 1849
537 ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe 556
1850 AGTTTTCATCATTTGACGCGGTTTCCAATCAGATTGAGCGCTGTAACTCTTTTAGCATTT 1909
557 GlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluGly 576

1910 GGAGGCTACTTAAGTAGTAGAAGAAATACTCCACCACAAATAATCGAAAAAGAGGC 1969
577 LeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu 596
1970 TTGGAGATAATGATTGGGAAGAAAGGCATCCAGGCTCATTAACAATCTTGTGTACTTA 2029
597 AspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArg 616
2030 GACTCAACCTTATTCTCTCTTATTGCTTTTAGTTCTGTCTGGACACTGTGTACTTAGA 2089
617 ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu 636
2090 CCCAAGAAAAAGAACGATGTAGATATATTAGTGAACCCCAAGAGCTACTGAGGACAGAA 2149
637 IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg 656
2150 ATTTGTTAATCCTCTGAGAAATATATGGATATGTGTGTGCCACAAAAAATATGAACCTGAGG 2209
657 LysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGlu 676
2210 AAAATACTTGAAGAGTGGAGGCTGCATCAGGATTTACCTCTGAAGAAAAAGATCCTGAG 2269
677 GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArg 696
2270 GAATTTCTTGAATATCTCTGTTTCATCATATTTTAAGGTTAGAACCTTTGCTTAAAAATAAGA 2329
697 SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu 716
2330 TCAGCAGGTCAAAAGGTACAGATTGTTACTTCTCATCAAAATTTTATCGAAAAAATGAG 2389
717 LysValGlyValProThrIleGlnGlnLeuLeuGluTrpSerPheIleAsnSerAsnLeu 736
2390 AAAAGTTGGCGTTCCCAATTCAGCAGTTGTTAGATGGTCTTTTATCAACAGATAACCTG 2449
737 LysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPhe 756
2450 AAAATTTGCAGAGGACCATCATGTCTGATTATTTCAGATGCTCGATTTGGAAAAAGACTTT 2509
757 LysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuGluAsp 776
2510 AAATATTAAAAAATTTTCTCTCTGGAATTAATAATACAGATTTTACTTGAAGAC 2569
777 ThrProArgGlnCysArgIleCysGlyLeuAlaMetTyrGluCysArgGluCysTyr 796
2570 ACTCCACAGACAGTCCCGATATGTGGAGGCTTGCATATGTATGAGTGTAGAGAAATGCTAC 2629
797 AspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGln 816
2630 GACGATCCGGACATCTCAGCTGGAAAAAATCAAGCAGTTTTTGTAAAAACCTGCAACACTCAA 2689
817 ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAsp 836
2690 GTCCACCTTCATCCGAAGAGGCTGAATCATTAATATATAACCCAGTGTCTACTTCCCAAGAC 2749
837 LeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAla 856
2750 TTACCCGACTGGGACTGGACACGCGTGCATCCCTTGGCCAGAAATATCGAGTTATTTCCT 2809
857 ValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSer 876
2810 GTTCTCTGCATAGAAAACAGCCACTATGTTGCTTTTGTGAAGTATGGGAAGACGATTTCT 2869
877 AlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIle 896
2870 GCCTGGCTCTTCTTTGACGATGGCGGATGGTGGTCAAGATGGGCTTCAACATT 2929
897 ProGlnValThrProCysProGluValGlyLysTyrLeuLysMetSerLeuGluAspLeu 916
2930 CCTCAAGTCAACCCATGCCAGAAAGTAGGAGTACTTTGAAGATGTCTCTCGAGAGACCTG 2989
917 HisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyr 936
2990 CATTCCTTGGACTCCAGAGAAATCCAAGGCTGTGCAGAGACTGCTTTGTGTGATCATAT 3049

QY 937 MetCysMetTyGlnSerProThrMetSerLeuTyLys 949
 DB 3050 ATGTGATGTACAGAGTCCAAACATGAGTTGTACAAA 3088

RESULT 12

US-09-434-737-1278
 ; Sequence 1278, Application US/09434737 — *Abdn*
 ; GENERAL INFORMATION:
 ; APPLICANT: Gearing, David P.
 ; APPLICANT: Kingsbury, Gillian A.
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
 ; TITLE OF INVENTION: HUMAN MIXED LYMPHOCYTE LIBRARY
 ; FILE REFERENCE: 1600.1067001
 ; CURRENT APPLICATION NUMBER: US/09/434,737
 ; CURRENT FILING DATE: 1999-11-05
 ; EARLIER APPLICATION NUMBER: 60/107,228
 ; EARLIER FILING DATE: 1998-11-05
 ; NUMBER OF SEQ ID NOS: 1830
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 1278
 ; LENGTH: 4668
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc.feature
 ; LOCATION: (1)...(4668)
 ; OTHER INFORMATION: n = A,T,C or G

US-09-434-737-1278

Alignment Scores:
 Pred. No.: 0 Length: 4668
 Score: 4983.00 Matches: 948
 Percent Similarity: 99.48% Conservative: 0
 Best Local Similarity: 99.48% Mismatches: 1
 Query Match: 98.99% Indels: 4
 DB: 22 Gaps: 4

US-09-671-687A-3 (1-949) x US-09-434-737-1278 (1-4668)

QY 1 MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTyTrpGluGluArgIle 20
 DB 230 ATGAGTTACGGCTATGGAGCCAAAGAAAGTCACTCCCTACTCGGGAAGACGGATT 289
 QY 21 PheTyLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLys 40
 DB 290 TTTTACTTCTCTTCAAGATCCAGCGTTACAGACAAACAAACAAAGCTCCCTTAA 349
 QY 41 ValProLysGlySerIleGlyGlnTyrlleGlnAspArgSerValGlyHisSerArgIle 60
 DB 350 GTACCGAAGGAGTATAGACAGTATATTCAAGATCGTTCTGTGGGCATTCAAGATT 409
 QY 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGlnProHisAla 80
 DB 410 CCTTCTGCAAAAGGCAAGAAATCAGATTGGATTAAATTAATAATCTAGAGCAACCTCATGCA 459
 QY 81 ValLeuPheValAspGlu--AspValValIleAsnGluLysPheThrGluLeu 99
 DB 470 GTTCTCTTTGTGATGAAAGGATGTTGTAGACATAAATGAAAGTTACAGAGTTACTT 529
 QY 100 LeuAlaIleThrAsnGlyGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
 DB 530 TTGGCAATTACCAATTGTGGAGAGAGTTTCAGGCTGTTTAAACACAGAACTAAGT 589
 QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
 DB 590 AAAGGCTCCAAATAGACGTGGGCTGTCTGTGAAAGTACAGCTGAGATCTGGGGAAGAA 649
 QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
 DB 650 AAATTTCTCGAGTTGTACGCTTTACAGAGGACCCCTGTTACAGAGAGGACAGTCTCCGGA 709
 QY 160 IlePhePheGlyValGluLeuLeuGluGluGlyArgGlyGlnGlyPheThrAspGlyVal 179

DB 710 ATATTTCTTTGGAGTTGAAATTCGTGGAGAAAGGTCGTGGTCAAGGTTTCACTGACGGGGTG 769
 QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
 DB 770 TACCAAGGGAACACAGCTTTTTCAGTGTGATGAAGATTGTGGCGTGTGTTGTCATTGGAC 829
 QY 199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyAlaGlyProGly 218
 DB 830 AAGCTAGAACTCATAGAGATGATGACATCTGCATTTGGAAAGTGATTACGCAAGGTCCTGGG 889
 QY 219 AspThrMetGlnValGluLeuProLeuGluIleAsnSerArgValSerLeuLysGly 238
 DB 890 GACACAAATGCAAGTCGAACCTTCTCTTTGGAATAAATCCAGAGTTTCTTTGAAGGTT 949
 QY 239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
 DB 950 GGAGAAACATAGAACTCGAACAGATTATTTCTGTGATGTTTTCAGGAGAAAGAAAGC 1009
 QY 259 LeuGlyTyPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe 278
 DB 1010 TTAGGATATTTTGTGTGGACATGGATAACCTATTGGCAACTGGGATGGAAGATT 1069
 QY 279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
 DB 1070 GATGGAGTCAGCTTTTGTAGTTTGTGCTGTGTTGAAAGTACAAATTTCTATTGCACATCAAT 1129
 QY 298 AspIleIleProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMetSer 317
 DB 1130 GATATCATCCAGAGAGTGTGCGAGGAAAGAGGCTCCCAAACTTGCCTTTATGTCA 1189
 QY 318 ArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThrGlySerThr 337
 DB 1190 AGAGGTGTTGGGACAAAGGTTTCATCCAGTCATAATAAACCAAGGCTACAGGATCTACC 1249
 QY 338 SerAspProGlyAsnArg---ArgSerGluLeuPheTyThrLeuAsnGlySerSerVal 356
 DB 1250 TCAGACCTCTGGAATAAGAAACAGATCTGAATATTTTATATACCTTAAATGGGCTCTCTGTT 1309
 QY 357 AspSerGlnProGlnSerLysSerLysAsnThrTrpTyrlleAspGluValAlaGluAsp 376
 DB 1310 GACTCAACCAACCAATCCAAATCAAAAATACATGGTACATTTGATGAAGTTGCAAGAAC 1369
 QY 377 ProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProProLeuGln 396
 DB 1370 CCTGCAAAATCTTTACAGAGATATCTACAGACTTTTCACCGTCTTCCACCACCTCCAG 1429
 QY 397 ProProValAsnSerLeuThrGluAsnArgPheHisSerLeuProPheSerLeu 416
 DB 1430 CCTCTCTCTGTAACCTCACTGACCCAGAACAGATTCCACTCTTTTACCATTCACTCTC 1489
 QY 417 ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGln 436
 DB 1490 ACCAAGATGCCCAATCCAAATGAAAGTATTTGGCCACAGTCCACTTCTCTGTACGCCAG 1549
 QY 437 SerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPro 456
 DB 1550 TCTGTATGGAAGAGCTAAACACTGCACCCGTCACAGAGAGTCCACCTTGGCCATGGCT 1609
 QY 457 ProGlyAsnSerHisGlyLeuGluValIleGlySerLeuAlaGluValLysGluAsnProPro 476
 DB 1610 CCTGGGAACCTCACATGGTCTAGAAGTGGGCTCATTTGGCTGAAGTTAAGGAGAACCTCTCT 1669
 QY 477 PheTyGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGly 496
 DB 1670 TTCTATGGGTAAATCCGTGTGGATCGGTGAGCCACAGGACTGAATGAAGTCTCGCTGGA 1729
 QY 497 LeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTy 516
 DB 1730 CTGAACTGGAAGATGATGTGACGGCTGTACGGATGGAACCTTCAGAGGACCTCGGTAT 1789
 QY 517 PheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSer 536

Db 1790 TTCACCTGTGCTGAAGAGCGCTGTTTGTGAACCTGAAGAGCTGCAGGCGCTGACTCT 1849
Qy 537 ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe 556
Db 1850 AGGTTTGCATCATTTGCAGCGGTTTCCATCATCAGATTGAGCGCTGTAACTCTTTAGCATTT 1909
Qy 557 GlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluGly 576
Db 1910 GGAGGCTACTTTAAGTGAAGTAGAGAGAAATACCTCCCAAAATGGAAAAAGAGGC 1969
Qy 577 LeuGluIleMetGlyLysLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu 596
Db 1970 TTGGAGATATGATTGGGAAGAGAGGATCCAGGGTCAATACAACTCTTGTACTTGA 2029
Qy 597 AspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArg 616
Db 2030 GACTCAACCTTATCTGCTTATTTGCTTTTGTGTTCTGACACTGTGTTACTTAGA 2089
Qy 617 ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu 636
Db 2090 CCCAAAGAAAGAACCATGTAGAAATATTATAGTGAACCCCAAGAGCTACTGAGACAGAA 2149
Qy 637 IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg 656
Db 2150 ATTGTTAATCCTTGAGATATATGATATGCTGTGCCACAAAATATTGAACCTGAGG 2209
Qy 657 LysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGlu 676
Db 2210 AAAATACTTTGAAAGGTGAGGCTGCATCAGGATTTTACCTCTCGAAGAAAAGATCTGAG 2269
Qy 677 GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArg 696
Db 2270 GAATCTCTGAATATCTGTTTCAATCATATATTTTAAGGGTAGAACCTTTGCTAAAAATAAGA 2329
Qy 697 SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu 716
Db 2330 TCAGCAGGTCAAAAGGTACAGGATTTGTTACTTCTCAAAATTTTATGGAAAAAATGAG 2389
Qy 717 LysValGlyValProThrIleGlnGlnLeuLeuGluTyrSerPheIleAsnSerAsnLeu 736
Db 2390 AAAGTTGGCTTCCCAATTCAGCAGTGTGTAGATGGTCTTTTATCAACAGTAACTTG 2449
Qy 737 LysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPhe 756
Db 2450 AAATTTGCAGAGGCACCATCATGCTGTGATATTATCAGATGCTCGATTTGGAAAAAGACTTT 2509
Qy 757 LysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAsp 776
Db 2510 AAACATTTTAAAAAAATTTTCTCTCTGGAATTTAAATATAACAGATTTTACTTGAAGAC 2569
Qy 777 ThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyr 796
Db 2570 ACTCCAGACAGTCCGGATATGTGAGGCTTGCATGTATGAGTGTAGAGATGCTAC 2629
Qy 797 AspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGln 816
Db 2630 GACGATCCGACATCTCAGCTGGAAAAATCAAGCAGTTTTTGTAAACCTCGCAACACTCAA 2689
Qy 817 ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAsp 836
Db 2690 GTCCACCTTCATCCGAAGAGGCTGAATATATAATATAAACCCAGTGTCACTTCCCAAGAC 2749
Qy 837 LeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAla 856
Db 2750 TTACCCGACTGGGACTGGAGACAGGCTGCATCCCTTGCAGAAATATGGAGTTATTGCT 2809
Qy 857 ValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSer 876
Db 2810 GTTCTCTGCATAGAAACAGCCACTATGTGCTTTTGTGAAGTATGGGAAGGAGGATCT 2869
Qy 877 AlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlnAsnGlyPheAsnIle 896
Db 2870 GCTGCTCTTCTTTGACAGCATGGCCGATCGGATGGTGGTTCAGAAATGGCTTCAACATT 2929

Qy 897 ProGlnValThrProCysProGluValGlyGlyTyrLeuLysMetSerLeuGluAspLeu 916
Db 2930 CCTCAAGTCACTCCCATGCCAGAGTAGGAGACTTGAAGATGTCTCTGGAAGACCTG 2989
Qy 917 HisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyr 936
Db 2990 CATTCCTGGACTCCAGGAGATCCAAAGGCTGTGCAGAGACTGCTTTGTGATGCATAT 3049
Qy 937 MetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
Db 3050 ATGTGCATGTACCAGAGTCCAAACAATGAGTTGTACAAA 3088

RESULT 13

US-09-850-118-1278 - *Abdn*
; Sequence 1278, Application US/09850118
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
; TITLE OF INVENTION: HUMAN MIXED LYMPHOCYTE LIBRARY
; FILE REFERENCE: 1600.1067-002
; CURRENT APPLICATION NUMBER: US/09/850.118
; PRIOR FILING DATE: 2001-05-07
; PRIOR APPLICATION NUMBER: US 60/107,228
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: US 09/434,737
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 1830
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1278
; LENGTH: 4668
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(4668)
; OTHER INFORMATION: n = A,T,C or G
US-09-850-118-1278

Alignment Scores:
Pred. No.: 0 Length: 4668
Score: 4983.00 Matches: 948
Percent Similarity: 99.48% Conservative: 0
Best Local Similarity: 99.48% Mismatches: 1
Query Match: 98.99% Indels: 4
DB: 36 Gaps: 4

US-09-671-687A-3 (1-949) x US-09-850-118-1278 (1-4668)

Qy 1 MetSerSerGlyLeuTyrSerGlnGluLysValThrSerProTyrTrpGluGluArgIle 20
Db 230 ATGAGTTCAAGCTTATGGAGCCCAAGAAAAGTCACTTCACTTCCCTACTGGAGAGCGGATT 289
Qy 21 PheTyrIleLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLys 40
Db 290 TTTTACTTGTCTTCTCAAGAAATGCAGCGTTACAGACAAACAAACACAAAAGCTCTCTAAA 349
Qy 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60
Db 350 GTACCGAAGGAAGATATAGGACAGTATATTCAAGATCGTTCTGTGGGCAATTCAGGATT 409
Qy 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla 80
Db 410 CCTTCTCTCAAAAGCAAGAAAATCAGATTGGATTAAAAATTTAGACCACTCATGCA 469
Qy 81 ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu 99
Db 470 GTTCTCTTTGTTGATGAAAGGATGTTGTAGAGATAAATGAAGATTTCACAGAGTTACTT 529
Qy 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
Db 530 TTGGCAATTACCAATTGTGAGGAGAGGTTTACGCTGTTTAAAAAACAAGAAACAGACTAAGT 589

QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
DB 590 AAAGCGCTCCAAATAGACGTGGGCTGCTCTGTGAAAGTACAGCTGAGATCTGGGGAGAA 649
QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
DB 650 AAATTTCTTGGAGTTGATGCTTTCAGAGACCCCTGTAGCAGAGGACAGTCTCCGGA 709
QY 160 IlePhePheGlyValGluLeuLeuGluGlyValArgGlyGlnGlyPheThrAspGlyVal 179
DB 710 ATATTTCTTGGAGTTGAAATGCTGTGGAAGGTCGTGCTCAAGTTTCTAGCAGGGGTG 769
QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
DB 770 TACCAGGGAAACAGCTTTTTCAGTGTGATGAAGATTGTGGCGTGTGTTGTCATTGGAC 829
QY 199 LysLeuLeuLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
DB 830 AAGCTAGAACTCATAGAAGATGATGACCTGCAATTGGAAGTGAATTACCGAGTCTCTGG 889
QY 219 AspThrMetGlnValGluLeuProLeuGluIleAsnSerArgValSerLeuLysGly 238
DB 890 GACACAATGCGAGTCTGAACTTCTCTTGGAAATPAACTCCAGAGTTTCTTTGAAGGTT 949
QY 239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
DB 950 CGAGAAACAATAGAACTCGAACAGTTATATCTGTGATGTTTTCGCCAGGAAAGAAGC 1009
QY 259 LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe 278
DB 1010 TTAGGATATTTGTTGGTGTGGACATGGATAACCTATTGGCAACTGGGATGGAAGATT 1069
QY 279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
DB 1070 GATGGAGTGCAGCTTTGTAGTTTGGTGTGTTGAAAGTACAATCTATTGGCACATCAAT 1129
QY 298 AspIleIleProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMetSer 317
DB 1130 GATATCATCCAGAGAGTGTGCGCAGGAAAGGAGGCTCCCAAACTTGCCTTTATGTCA 1189
QY 318 ArgGlyValGlyAspLysGlySerSerSerHisAsnLysProLysAlaThrGlySerThr 337
DB 1190 AGAGGTGTTGGGACAAAGGTTTCATCCAGTCAATAAACAAGGCTACAGGATCTACC 1249
QY 338 SerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVal 356
DB 1250 TCAGACCTCGGAAATAGAAACAGATCTGAATTAATTTTATACCTTAAATGGGCTCTCTGTT 1309
QY 357 AspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGluValAlaGluAsp 376
DB 1310 GACTCACACCAATCCAAATCAAAAATACATGTTGATGAGTTGCGAGAGAC 1369
QY 377 ProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerProProLeuGln 396
DB 1370 CCTGCAAAATCTCTTACAGAGATATCTACAGACTTTTGACCGTCTTCCACCACCTCCAG 1429
QY 397 ProProProValAsnSerIleThrThrGluAsnArgPheHisSerLeuProPheSerLeu 416
DB 1430 CCTCTCTGTGAACCTCACTGACCCGAGAACAGATTCACCTCTTTTACCATTTCAGTCTC 1489
QY 417 ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGln 436
DB 1490 ACCAAGATGCCCAATCAATGGAAGTATTGGCCACAGTCCACTTCTCTGTGAGCCAG 1549
QY 437 SerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPro 456
DB 1550 TCTGTAATGGAAGAGCTAAACACTGCACCGCTCCAGAGAGTCCACCCTTTGGCCATGCTC 1609
QY 457 ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro 476
DB 1610 CCTGGGAACCTCACATGCTTAGAAGTGGGCTCATTTGCTGAAGTTAAGGAGAACCCCTCT 1669

QY 477 PheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGly 496
DB 1670 TTCTATGGGGTAAATCCGTTGGATCGGTCCGCCACAGGACTGAATGAAGTCTCGCTGGA 1729
QY 497 LeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr 516
DB 1730 CTGGAACTGGAAGATGAGTGTGCGGCTGTACGATGGAAACCTTCAGAGGACATCGGTAT 1789
QY 517 PheThrCysAlaLeuLeuLysAlaLeuPheValLysLeuLysSerCysArgProAspSer 536
DB 1790 TTCACCTGTGCTGGAAGAGCGCTGTTGTGAACCTGAAGAGCTGCAGGCTGACTCT 1849
QY 537 ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe 556
DB 1850 AGGTTTGCATCATTTGCAGCGGTTTCCAATCAGATTGAGCGCTGTAACTCTTTAGCATTT 1909
QY 557 GlyGlyTyrLeuSerGluValValGluAsnThrProProLysMetGluLysGluGly 576
DB 1910 GGAGGCTACTTAAGTGAAGTAGAAGAAAAATACTCCACAAAAATGGAAGAAAGAGGC 1969
QY 577 LeuGluIleMetIleGlyLysLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu 596
DB 1970 TTGGAGATTAATGATTGGGAAGAAAGGCATCCAGGTCATTACAACTCTTTACTTA 2029
QY 597 AspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArg 616
DB 2030 GACTCAACCTTATCTGCTTATTTGCTTTTAGTTCTGTTCTGGACACTGTGTTACTTAGA 2089
QY 617 ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu 636
DB 2090 CCCAAAGAAAAGAACCATGTAGAAATATTATAGTGAACCCCAAGAGCTACTGAGGACAGAA 2149
QY 637 IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg 656
DB 2150 ATGTGTAATCTCTGAGAATATATGATGATGTGTGTGCCACAAAAATTTAGAACTCAGG 2209
QY 657 LysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGlu 676
DB 2210 AAAATACTTTGAAAGGTGGAGGCTGCATCAGGATTTACCTCTGAAGAAAAAAGATCTCTGAG 2269
QY 677 GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArg 696
DB 2270 GAATTTCTTGAATATCTGTTTTCATCATATTTTAAAGGTTAGAACCTTTTGTCTAAATAAGA 2329
QY 697 SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu 716
DB 2330 TCAGCAGGTCAAAGGTACAGATTTGTTACTCTTCAAAATTTTTATGGAAAAAATGAG 2389
QY 717 LysValGlyValProThrIleGlnGlnLeuLeuGluTrpSerPheIleAsnSerAsnLeu 736
DB 2390 AAAGTTGGGCTTCCACATTCAGCAGTTGTTAGNATGGTCTTTTATCAACAGTAACCTG 2449
QY 737 LysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPhe 756
DB 2450 AAATTTGCAGAGGACCATCATGCTCTGATTTATTCAGATGCTCGATTTGGAAAAAGACTTT 2509
QY 757 LysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAsp 776
DB 2510 AAATCTTTTAAAAAAATTTTTCTCTCTGGAATTAATAATATAACAGATTTACTTTGAAGAC 2569
QY 777 ThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyr 796
DB 2570 ACTCCACAGACAGTCCGGATATGTGGAGGCTTGGCAATGTATGATGTAGAGATGCTAC 2629
QY 797 AspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGln 816
DB 2630 GACGATCCGACATCTCAGCTGGAAAAATCAAGCAGTTTTTGTAAAACTCTGCAACACTCAA 2689
QY 817 ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAsp 836
DB 2690 GTCCACCTTCATCCGAGAGGCTGAATCATAAATATAAACCCAGGTGTCTCCTTCCCAAGAC 2749
QY 837 LeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAla 856

Db 2750 TTACCGACTGGAGCTGGAGACACGGCTGCATCCCTTGCAGATATGAGTTATTGCT 2809
QY 857 ValLeuCyLeuThrSerHisTyrValAlaPheValLysTyrGlyLysAspSer 876
Db 2810 GTTCTCTGCATAGAAACAAGCCACTATGTTGCTTTGTGAAGTATGGGAAGGACGATTCT 2869
QY 877 AlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyClnAsnGlyPheAsnIle 896
Db 2870 GCGTGGCTCTTTTACAGCATGGCCGATGGGATGGTGGTGCAGATGGCTTCAACATT 2929
QY 897 ProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLeu 916
Db 2930 CTTCAAGTCAACCCCATGCCAGAGTAGAGAGTACTTGAAGATGCTCTGGGAAGACCTG 2989
QY 917 HisSerLeuAspSerArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyr 936
Db 2990 CATTCCTGGACTCCAGGAGATCCAAAGCTGTGCAGGAGACTGCTTGTGATGCATAT 3049
QY 937 MetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
Db 3050 ATGTGCATGTACAGAGTCCAAATGAGTTGTACAAA 3088

RESULT 14

PCT-US02-27777-49
; Sequence 49, Application PC/TUS0227777
; GENERAL INFORMATION:
; APPLICANT: diaDexus, Inc.
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; APPLICANT: Salceda, Susana
; TITLE OF INVENTION: Compositions and Methods Relating to Breast Specific Genes and P
; FILE REFERENCE: DEX-0346
; CURRENT APPLICATION NUMBER: PCT/US02/27777
; CURRENT FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: US 60/316,306
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 49
; LENGTH: 6831

} too late

; TYPE: DNA

; ORGANISM: Homo sapien

PCT-US02-27777-49

Alignment Scores:
Pred. No.: 0 Length: 6831
Score: 4983.00 Matches: 948
Percent Similarity: 99.48% Conservative: 0
Best Local Similarity: 99.48% Mismatches: 1
Query Match: 98.99% Indels: 4
DB: 1 Gaps: 4

US-09-671-687A-3 (1-949) x PCT-US02-27777-49 (1-6831)

QY 1 MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluGluArgIle 20
Db 1585 ATGAGTTTCAGGCTTATGGAGCCCAAGAAAAGTCACTTCCCTTACTGGGAAGAGCGATT 1644
QY 21 PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLys 40
Db 1645 TTTTACTTCTCTTCAAGATGCAGGCTGTACAGACAAACAAACAAAGCTCCTTAAA 1704
QY 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60
Db 1705 GTACCGAAGGGAAGTATAGACAGTATATTCAAGATCGTCTCTGTGGGCAATCAAGATT 1764
QY 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGlnProHisAla 80
Db 1765 CCTTCTGCAAAAGGCAAGAAAATCAGATTGGATTAAATAATTCTAGACCAACCTCATGCA 1824
QY 81 ValLeuPheValAspGlu---AspValValGluLeuAsnGluLysPheThrGluLeu 99

Db 1825 GTTCTCTTTTGTGATGAAGAAGGATGTTGTAGAGATAAATGAAGAAGTTCCACAGAGTTACTT 1884
QY 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
Db 1885 TTGGCAATTTACCAATTTGTGAGGAGAGGTTTCCAGCTGTTTAAAAACAGAAAACAGACTAAGT 1944
QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
Db 1945 AAAGGCTTCCAAATAGACGTGGGCTGCTGTGTAAGTACAGCTGAGATCTGGGGAAGAA 2004
QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
Db 2005 AAATTTCTGGAGTTGTACGCTTTCAGAGACCCCTGTTAGCAGAGAGACAGTCTCCGGA 2064
QY 160 IlePhePheGlyValGluLeuLeuGluGluGlyArgGlyGlnGlyPheThrAspGlyVal 179
Db 2065 ATATTCTTGGAGTTGAATTTGCTGGAAGAAGTCTGCTGCTCAAGGTTTTCACCTGACGGGGTG 2124
QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
Db 2125 TACCAAGGGAACACGCTTTTTCAGTGTGATGAAGATTGTGGCGGTGTTGTTGCAATTGGAC 2184
QY 199 LysLeuGlnLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
Db 2185 AAGCTAGAACTCATAGAAATGATGACATGCACTTGGAAAGTATTACCGAGGCTCTGGG 2244
QY 219 AspThrMetGlnValGluLeuProLeuGluLeuAsnSerArgValSerLeuLysGly 238
Db 2245 GACACAATGCAGGTCGAACCTTCTCTTGGAAATAAATCCACAGAGTTTCTTTGAAGGTT 2304
QY 239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
Db 2305 GGAGAAACAATAGAAATCTGGAACAGATTATATCTGTGATGTTTTCGCCAGGAAAAAGAACG 2364
QY 259 LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe 278
Db 2365 TTAGGATATTTTGTGTGGTGGACATGATAACCTATTGGCAACTGGGATGGGAAGATT 2424
QY 279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
Db 2425 GATGGAGTGCAGCTTTGTAGTTTGTGTGTTGAAAGTACAATTTCTATTGCACATCAAT 2484
QY 298 AspIleIleProGluSerValThrGlnGluArgProProLysLeuAlaPheMetSer 317
Db 2485 GATATCATCCAGAGAGTGTGACGAGAAAGGAGGCTCCCAAACTTGCCTTTATGTCA 2544
QY 318 ArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThrGlySerThr 337
Db 2545 AGAGGTGTTGGGGACAAAGGTTTCATCCAGTCATAATAAACCAAGGCTTACAGGATCTACC 2604
QY 338 SerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVal 356
Db 2605 TCAGACCTTGGAAATAGAAACAGATCTGAATTTATTTATACCTTAAATGGGTCTTCTGTT 2664
QY 357 AspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGluValAlaGluAsp 376
Db 2665 GATTCACACACCAATCCAAATCAAAAAATACATGTTGATGATGATGATGATGATGATGAT 2724
QY 377 ProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProProLeuGln 396
Db 2725 CCTGCAAAATCTCTTACAGAGATATCTACAGACTTTTCACCGTTTTCACCACTCCAG 2784
QY 397 ProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeu 416
Db 2785 CCTCTCTCTGTGAACCTCACTGACCAACAGAGATTTCACCTCTTTTACCACTTTCAGTCTC 2844
QY 417 ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGln 436
Db 2845 ACCAAGATGCCCAATCAATGGAAGTATTGGCCACAGTCCACCTTCTCTGTGACCCAG 2904
QY 437 SerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPro 456
Db 2905 TCTGTAAATGAAGAGCTAAACACTGCAACCCCGTCCCAAGAGAGTCCACCTTGGCCATGCT 2964

```
QY 457 ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro 476
Db 2965 CTGGGAACCTCAGATGCTAGAAAGTGGGCTCATTTGGCTGAAGTTAAGGAGAACCCCTCCT 3024
QY 477 PheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGly 496
Db 3025 TTCATATGGGGTAAATCCGTTGGATCGGTGAGCCAGCCAGGACTGAATGAAGTGGCTCGCTGGA 3084
QY 497 LeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr 516
Db 3085 CTGGAACTGGAAAGTGAAGTGGAGCTGACGGATGGAACCTTCAGAGGCACTCGGTAT 3144
QY 517 PheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSer 536
Db 3145 TTCACCTGTGCCCTGAAGAAGGCGCTGTTGTGAACCTGAAGAGCTGCAGGCGCTGACTCT 3204
QY 537 ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe 556
Db 3205 AGCTTTGCAATCATTTGAGCGCGGTTCCAAATCAGATTGAGCGCTGTAACTCTTTAGCATTT 3264
QY 557 GlyGlyTyrLeuSerGluValGluGluAsnThrProProLysMetGluLysGluGly 576
Db 3265 GGAGGCTACTTAAGTGAAGTAGTAGAAGAAATACTCCACCAAAATGGAAAAAGAGGC 3324
QY 577 LeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu 596
Db 3325 TTGGAGATAATGATTGGGAAGAGAAAGGCATCCAGGGTCATTACAAATCTTGTACTTA 3384
QY 597 AspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArg 616
Db 3385 GACTCAACCTTATCTGCTTATTTGCTTTTGTAGTTCTGTTCTGACACTGTGTTACTTAGA 3444
QY 617 ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnLeuLeuArgThrGlu 636
Db 3445 CCCAAAGAAAAGAACCATGTAGAAATATTATAGTGAACCCAGAGCTACTGAGACAGAA 3504
QY 637 IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg 656
Db 3505 ATTGTTAATCTCTGAGATATATGATGATGCTGTGCCACAAAATTTATGAACTGAGG 3564
QY 657 LysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGlu 676
Db 3565 AAAATACITTTGAAAGGTGGAGGCTGCATCAGGATTTTACCTCTCAAGAAAAGATCCTGAG 3624
QY 677 GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLysIleArg 696
Db 3625 GAATCTTGAATATCTGTTTTCATCATATTATTTAAGGGTAGAACCTTTGTCTAAAATAAGA 3684
QY 697 SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu 716
Db 3685 TCAGCAGGTCAAAAGGTACAAAGATTGTTACTTCTATCAAAATTTTATGGAAAAAATGAG 3744
QY 717 LysValGlyValProThrIleGlnLeuLeuGluTrpSerPheIleAsnSerAsnLeu 736
Db 3745 AAAGTTGGCGTTCACCAATTCAGCAGTGTGTAAGATGGTCTTTTATCAACAGCTAACCTG 3804
QY 737 LysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPhe 756
Db 3805 AAATTTGACAGAGGCACCATCATCTCTGATTATTTACAGATGCTCGATTTGGAAAAAGACTTT 3864
QY 757 LysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAsp 776
Db 3865 AAATTAATTAATAAATTTTCTCTCTGGAATTAATAATAACAGATTTACTTTGAAGAC 3924
QY 777 ThrProArgGlnCysArgIleCysGlyLeuAlaMetTyrGluCysArgGluCysTyr 796
Db 3925 ACTCCAGACAGTCCGGATATGTGGAGGCTTGCATGATGATGATGATGATGATGATGATGAT 3984
QY 797 AspAspProAspIleSerAlaGlyLysIleGlnPheCysValThrCysAsnThrGln 816
Db 3985 GACGATCCGGACATCTCAGCTGGAAAAAATCAAGCAGTTTTTGTAAAAACCTGCAACACTCAA 4044
```

```
QY 817 ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAsp 836
Db 4045 GTCCACCTTTCTCCGAAGAGGCTGAATATAAATATAAACCAGTGTCACTTCCCAAAGAC 4104
QY 837 LeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAla 856
Db 4105 TTACCCGACTGGAGCTGGAGACACGGCTGCATCCCTTCCAGAAATATGGAGTTATTGCT 4164
QY 857 ValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSer 876
Db 4165 GTTCTCTGCATAGAAACAACCCACTATGTTGCTTTTGTGAAGTATGGAAGGACGATTCT 4224
QY 877 AlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIle 896
Db 4225 GCCTGCTCTCTTTTACAGCATGGCCGATCGGATCGGATCGGATCGGATCGGATCGGAT 4284
QY 897 ProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLeu 916
Db 4285 CCTCAAGTCACCCCATGCCAGAAAGTAGGAGAGTACTTTGAAGATGTCCTCGGAAGACCTG 4344
QY 917 HisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyr 936
Db 4345 CATTCTCTGGACTCCAGAGAAATCCAAAGCTGTGCAGAGAGACTGCTTTGTGATGCATAT 4404
QY 937 MetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
Db 4405 ATGTGCATGTACCAGAGTCCAAACATGAGTTTGTACAAA 4443

RESULT 15
PCT-US02-27777A-49
; Sequence 49, Application PC/TUS0227777A
; GENERAL INFORMATION:
; APPLICANT: diaDexus, Inc.
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; APPLICANT: Salceda, Susana
; TITLE OF INVENTION: Compositions and Methods related to Breast Specific Genes and Prot
; FILE REFERENCE: DEX-0346
; CURRENT APPLICATION NUMBER: PCT/US02/27777A
; PRIOR FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: US 60/316,307
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 49
; LENGTH: 6831
; TYPE: DNA
; ORGANISM: Homo sapien
PCT-US02-27777A-49

Alignment Scores:
Pred. No.: 0 Length: 6831
Score: 4993.00 Matches: 948
Percent Similarity: 99.48% Conservative: 0
Best Local Similarity: 99.48% Mismatches: 1
Query Match: 98.99% Indels: 4
DB: 1 Gaps: 4

US-09-671-687A-3 (1-949) x PCT-US02-27777A-49 (1-6831)
QY 1 MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluGluArgIle 20
Db 1585 ATGAGTTTCAGCTTATGGAGCCCAAGAAAGTCACTTCACTCCCTACTCGGAAGACGGATT 1644
QY 21 PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLys 40
Db 1645 TTTTACTTGTCTTCTCAAGAAATGCAGCGTTACACACAAACAAACACAAAGCTCCTTAAA 1704
QY 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60
Db 1705 GTACCCGAAGGGAAGTAGTAGACAGTATATTCAAGATCGTTCTGTGGGCGATTCAAGGATT 1764
QY 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla 80
```

} too late

1765 CTTTCTGCAAAAGGCAAGAAAATCAGATTGGATTAAATAATCTAGAGCAACCTCATGCA 1824
81 ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu 99
1825 GTTCTCTTTGTTGATGAAAGGATGTTGATAGATAAATGAAAGTTTCACAGAGTTACTT 1884
100 LeuAlaIleThrAsnGluLeuPheSerLeuPheLysAsnArgAsnArgLeuSer 119
1885 TTGGCAATTTACCAATTTGTCAGGAGAGTTGAGCTGTTTAAACACAGAACACAGACTAAGT 1944
120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
1945 AAAGGCTCCAAATAGACGTGGCTGTCTGTGAAAGTACAGCTGAGATCTGGGGAGAA 2004
140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
2005 AAATTTCTTGGAGTTGTACGCTTCAGAGGACCCCTGTAGCAGAGGACAGTCTCCGGA 2064
160 IlePhePheGlyValGluLeuLeuGluGluGlyArgGlyGlnGlyPheThrAspGlyVal 179
2065 ATATTCTTTGGAGTTGAAATGCTGGAGAGGTCGTGGTCAAGGTTTCACTGACGGGGTG 2124
180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
2125 TACCAAGGAAACAGCTTTTTCAGTGTGATGAAGATTGTGGCGTGTGTTGTCATTGGAC 2184
199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
2185 AAGCTAGAACTCATAGAAATGATGACACTGCTGTAAGAGTGATTTACGAGGTCCTGGG 2244
219 AspThrMetGlnValGluLeuProLeuGluIleAsnSerArgValSerLeuLysGly 238
2245 GACACAAATCGAGTCGAACTTCCTCTTTGGAAATAAACTCCAGAGTTTCTTTGAAAGTT 2304
239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
2305 GGAGAAACAATAGAACTCGAAACAGTTATATCTGTGATGTTTGGCAGGAAAGAAAGC 2364
259 LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe 278
2365 TTAGGATATTTGTTGGTGGACATGGATTAACCTATTGGCAACTGGGATGGAAAGATTT 2424
279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
2425 GATGGAGTCAGCTTTGTAGTTTGGTGTGTGAAAGTACAATTTCTATTGCAATCAAT 2484
298 AspIleIleProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMetSer 317
2485 GATATCATCCAGAGAGTGTGACGAGGAAAGGAGGCTCCCAAACTTGCCTTTATGTCA 2544
318 ArgGlyValGlyAspLysGlySerSerSerHisAsnLysProLysAlaThrGlySerThr 337
2545 AGAGGTGTTGGGACAAAGGTTTCATCCAGTCATAATAAACCAAGGCTACAGATCTACC 2604
338 SerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVal 356
2605 TCAGACCCCTGGAAATAGAAACAGATCTGAATTAATTTTATACCTTAATGGGCTTCTGTT 2664
357 AspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGluValAlaGluAsp 376
2665 GACTCACACCAATCCAAATCAAAAATACATGTTGATGTAAGTTGAGAGAGAC 2724
377 ProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProLeuGln 396
2725 CTTGCAAAATCTTTACAGAGATATCTACAGACTTTTGACCGTTCCTTCCACCACCTCCAG 2784
397 ProProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeu 416
2785 CCTCTCTCTGTGAACACTCACTGACCCAGGAGAACAGATTCCACTCTTTTACCATTCAGTCTC 2844
417 ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGln 436

2845 ACCAAGATGCCCAATACCAATGGAAGTATTGGCCACAGTCCACTTTCTCTGTGAGCCAG 2904
437 SerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPro 456
2905 TCTGTAATGGAAGAGCTAAACACTGCACCCGTCACAGAGAGTCCACCCTTGCCCATGCT 2964
457 ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro 476
2965 CCTGGGAACCTCACATGCTTAGAAGTGGCTCATTTGGCTGAAGTTAAAGAGAACCCCTCT 3024
477 PheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGly 496
3025 TTCTATGGGTGTAATCCGTTCGATCGGTGAGCCACAGGACTGAATGAAGTGTCTCGTGA 3084
497 LeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr 516
3085 CTGGAACCTGGAAGATGATGTGACAGGCTGTACGATGGAACCTTCAGAGGACATCGGTAT 3144
517 PheThrCysAlaLeuLysLysAlaLeuPheValLysLysSerCysArgProAspSer 536
3145 TTCACCTGTCCCTGAAAGAGGCGCTGTTGTGAACTGAAAGAGCTGCAGGCCCTGACTCT 3204
537 ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe 556
3205 AGTTTGCATCATTGACCGGTTTCCATCAGATTGAGCGCTGTAACTCTTTAGCATTT 3264
557 GlyGlyTyrLeuSerGluValValGluLeuAsnThrProProLysMetGluLysGluGly 576
3265 GGAGGCTACTTAAGTGAAGTAGTAGAAGAAATACTCCACCACCAAAATGAAAAAGAGGC 3324
577 LeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu 596
3325 TTGAGATAATGATTGGGAAGAGAAAGGCATCCAGGCTCATTAACAATCTTGTGTACTTA 3384
597 AspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArg 616
3385 GACTCAACCTTATCTGCTTATTTGCTTTAGTCTGTGTCGACACTGTGTACTTAGA 3444
617 ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu 636
3445 CCCAAGAAAAAGACGATGTAGATATATAGTGAACCCCAAGAGCTACTGAGGACAGNA 3504
637 IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg 656
3505 ATTGTTAATCTCTGAGATATATGATGTGTGTGTCACAAAAATTTATGAAACTGAGG 3564
657 LysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGlu 676
3565 AAAATATCTTGAAAGGTGGAGGCTGCATCAGGATTTACCTCTGAAGAAAAAGATCCTGAG 3624
677 GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArg 696
3625 GAATCTTGAATATCTGTTTCATCATATTTAAGGGTAGAACCTTTCTTAAAAATAGA 3684
697 SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu 716
3685 TCAGCAGGTCAAAAGGTACAGATTTGTTACTTCTTCAATTTTATCGAAAAAATGAG 3744
717 LysValGlyValProThrIleGlnGlnLeuLeuGluTrpSerPheIleAsnSerAsnLeu 736
3745 AAAGTTGGCGTTCCCAATTCAGCAGTTGTAGATGGTCTTTTATCAACAGATAACCTG 3804
737 LysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPhe 756
3805 AAATTTGTCAGAGGACCATCATGTCTGATTTTACAGATGCGCTCGGATTTGGAAAAAGACTTT 3864
757 LysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAsp 776
3865 AAATATTTAAAAAATTTTCTCTCTGGAATTAATATAACAGATTTTACTTGAAGAC 3924
777 ThrProArgGlnCysArgIleCysGlyLeuAlaMetTyrGluCysArgGluCysTyr 796
3925 ACTCCACAGACAGTCCCGGATATGTGGAGGCTTCGAATGTATGATGATAGAGAAATGCTAC 3984

```
QY 797 AspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGln 816
Db |||||||
3985 GACGATCCCGACATCTCAGCTGGAAAAATCAAGCAGTTTTGTAAAAACCTGCAACACTCAA 4044
QY 817 ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAsp 836
Db |||||||
4045 GTCCACCTTCATCCGAAAGGCTGAATCATATAATATAACCCAGTGTCACCTCCCAAAGAC 4104
QY 837 LeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAla 856
Db |||||||
4105 TTACCCGACTGGGACTGGAGACACGCGCTGCATCCCTTGCAGAAATATGGAGTTATTTGCT 4164
QY 857 ValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSer 876
Db |||||||
4165 GTTCTCTGCATAGAAACAAGCCACTATGTTGCTTTTGTGAAGTATGGGAAGGACGATTCT 4224
QY 877 AlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIle 896
Db |||||||
4225 GCGTGGCTCTTCTTTGACAGCATGGCCGATCGGGATGGTGGTCAGAAATGGCTTCAACATT 4284
QY 897 ProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLeu 916
Db |||||||
4285 CCTCAAGTCACCCATGCCAGAGTAGAGAGTACTTGAAGATGCTCTCTGGAGACCTG 4344
QY 917 HisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyr 936
Db |||||||
4345 CATTCTTGGACTCCAGGAGATCCAAGGCTGTGCACGAAGACTGCTTTGTGATGCATAT 4404
QY 937 MetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
Db |||||||
4405 ATGTGCATGTACCAGAGTCCACAATGAGTTGTACAAA 4443
```

Search completed: April 18, 2005, 22:14:04
Job time : 7268 secs